

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 18:00:14 ; Search time 8809 Seconds  
(without alignments)  
11257.675 Million cell updates/sec

Title: US-10-009-791-21  
Perfect score: 2288  
Sequence: 1 gcaagagccacacgattaca.....acattctcaattttgtgaat 2288

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*  
29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pin.\*  
35: em.htg.rod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2288	100.0	2288	6	AX046795 Sequence
2	805.8	35.2	2291	8	AF036340 Arabidops
3	805.8	35.2	2335	8	AY045625 Arabidops
4	804.6	35.2	1779	8	AY133556 Arabidops
5	685.8	30.0	2088	8	AY168645 Oryza sat
6	683.2	29.9	711	8	AF536527 Glycine m
7	683	29.9	2163	8	AK121543 Oryza sat
8	668.4	29.2	2240	6	AX046793 Sequence
9	666.8	29.1	2305	8	AK101514 Oryza sat
10	581.6	25.4	2300	8	AK066392 Oryza sat
11	581.6	25.4	2924	8	AK100694 Oryza sat
12	428	18.7	108847	8	AF002109 Arabidops
13	412	18.0	482	6	AX046779 Sequence
14	376.8	16.5	134982	8	AP003279 Oryza sat
15	372.4	16.3	1728	6	AX046791 Sequence
16	362.6	15.8	128532	8	AC130602 Oryza sat
17	362.6	15.8	138467	2	AC130612 Oryza sat
18	300.8	13.1	115270	8	AC135205 Oryza sat
19	296.6	13.0	844	6	AX046777 Sequence
20	233.2	10.2	1074	6	AX046801 Sequence
21	233.2	10.2	1074	8	BT009618 Triticum
22	223.4	9.8	794	6	AX046781 Sequence
23	131.6	5.8	1392	6	AX660749 Sequence
24	129	5.6	701	6	AX046775 Sequence
25	119.4	5.2	2646	8	AF139835 Populus t
26	114.8	5.0	562	6	AX046789 Sequence
27	113	4.9	2147	8	AK100862 Oryza sat
28	111.4	4.9	2337	8	AF327430 Arabidops
29	110.6	4.8	134976	8	AF004968 Lotus cor
30	110.4	4.8	1816	8	BT001946 Arabidops
31	110.4	4.8	2225	8	AF005048 Arabidops
32	107.8	4.7	486	6	AX046799 Sequence
33	105.8	4.6	2956	8	AK072338 Oryza sat
34	105.2	4.6	266	6	AR246507 Sequence
35	97.2	4.2	139855	8	OSJN00054
36	96.6	4.2	1872	8	BT006340 Arabidops
37	96.6	4.2	2554	8	AY091772 Arabidops
38	96.2	4.2	1789	8	AY150427 Arabidops
39	96.2	4.2	2070	8	AF291816 Arabidops
40	94.6	4.1	2052	8	AY045799 Arabidops
41	94.4	4.1	140234	8	AC016779 Genomic s
42	94.4	4.1	164899	8	AC084218 Oryza sat
43	92.8	4.1	2123	8	AK100506 Oryza sat
44	92.8	4.1	2609	8	AK111900 Oryza sat
45	92.2	4.0	1801	8	BT002118 Arabidops

ALIGNMENTS

RESULT 1	AX046795	2288 bp	DNA	linear	PAT 15-DEC-2000
LOCUS	Sequence 21 from Patent WO0068406.				
DEFINITION	AX046795				
ACCESSION	AX046795.1	GI:11876306			
VERSION					
KEYWORDS	Glycine max (soybean)				
SOURCE	Glycine max				
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.				
REFERENCE	1				
AUTHORS	Caimi, P.G., Famodu, O.O., Lee, J.M., Miao, G.H. and Maxwell, C.A.				

TITLE		Disease resistance factors	
JOURNAL		Patent: WO 0068406-A 21 16-NOV-2000;	
FEATURES		E.I. DU PONT DE NEMOURS AND COMPANY (US)	
source	Location/Qualifiers		
	1..2288		
	/organism="Glycine max" /mol_type="unassigned DNA" /db_xref="taxon:3847"		
ORIGIN			
Query Match		100.0%; Score 2288; DB 6; Length 2288;	
Best Local Similarity		100.0%; Pred. No. 0;	
Matches 2288; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 GCACGAGGCGACACGTTTACACAGGCGACATATGGTTGCCGGAACAAATCCGATGGGAA 60		
Db	1 GCACGAGGCGACACGTTTACACAGGCGACATATGGTTGCCGGAACAAATCCGATGGGAA 60		
QY	61 GGGTGATATAGTGTCTTAGAGTAATATTTGTGTATTAACAGAACCGCGTTTGAAGCAGT 120		
Db	61 GGGTGATATAGTGTCTTAGAGTAATATTTGTGTATTAACAGAACCGCGTTTGAAGCAGT 120		
QY	121 GACGTGTATACAGTACATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 180		
Db	121 GACGTGTATACAGTACATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 180		
QY	181 AGTTTGTGCTTTACACACCATCTGTTTGGCCCTACCATTTCTCACTCATCATCATCATCAT 240		
Db	181 AGTTTGTGCTTTACACACCATCTGTTTGGCCCTACCATTTCTCACTCATCATCATCATCAT 240		
QY	241 AACCCATTCCTTTTGGCCCTTGAACAAACCTCTGCACCTTTTCTTTTCACTCTCAG 300		
Db	241 AACCCATTCCTTTTGGCCCTTGAACAAACCTCTGCACCTTTTCTTTTCACTCTCAG 300		
QY	301 TCTCCGATCCATATGACGAGGAAACGACGTCGCGGACACACGTCGTGTCAGCTGTGT 360		
Db	301 TCTCCGATCCATATGACGAGGAAACGACGTCGCGGACACACGTCGTGTCAGCTGTGT 360		
QY	361 CTTGAGTGGTCTATCCCTTACATGACGACCCCAAGGACCGGACCGGTTTCCACAGGT 420		
Db	361 CTTGAGTGGTCTATCCCTTACATGACGACCCCAAGGACCGGACCGGTTTCCACAGGT 420		
QY	421 GTGTGACGCTGTGTAAGAGTCTGCTGCTCAACCGGACGACGTCACCATCGCGCTCTG 480		
Db	421 GTGTGACGCTGTGTAAGAGTCTGCTGCTCAACCGGACGACGTCACCATCGCGCTCTG 480		
QY	481 CTACACCAACCCCGGCTCGCTCCGCGCGCTTCCGCGCTCGAGTCGCTCAAGCT 540		
Db	481 CTACACCAACCCCGGCTCGCTCCGCGCGCTTCCGCGCTCGAGTCGCTCAAGCT 540		
QY	541 CAAGGCGAAGCCCGGACCGCAATGTTTCACTTTGATACCGGAGTTGGGCGGACACGT 600		
Db	541 CAAGGCGAAGCCCGGACCGCAATGTTTCACTTTGATACCGGAGTTGGGCGGACACGT 600		
QY	601 CACTCCCTGGGTCAAGAGATTTCTCAGTACTTTCGATTGCTCAAGAGCTTCCACTCCG 660		
Db	601 CACTCCCTGGGTCAAGAGATTTCTCAGTACTTTCGATTGCTCAAGAGCTTCCACTCCG 660		
QY	661 CCGCATGATTGTCAAGGATTTCCGATCTTCAGATCTCGCTCGTGACCGCGGTCACTGCT 720		
Db	661 CCGCATGATTGTCAAGGATTTCCGATCTTCAGATCTCGCTCGTGACCGCGGTCACTGCT 720		
QY	721 TCAGGCTCTCAAGCTTGACAGTCTCCGCTTTCACACCGATGCTTTTTCATATCGG 780		
Db	721 TCAGGCTCTCAAGCTTGACAGTCTCCGCTTTCACACCGATGCTTTTTCATATCGG 780		
QY	781 TCGCTTTTGCAGAGTTTAAAGTCTTGTGTTTGGAGGAAGCTCAATTTCTGAGAGGA 840		
Db	781 TCGCTTTTGCAGAGTTTAAAGTCTTGTGTTTGGAGGAAGCTCAATTTCTGAGAGGA 840		
QY	841 CGGAGATGGCTACACGAGCTTCTTGAATAATACAGTCTTTCAGAGCTCTCAATTTT 900		
Db	841 CGGAGATGGCTACACGAGCTTCTTGAATAATACAGTCTTTCAGAGCTCTCAATTTT 900		
QY	901 CTTGACAGACATGCTGTGTGAAGATTGAGACCTTGAACTTTTAGCTTAAATAATTGCC 960		
Db	901 CTTGACAGACATGCTGTGTGAAGATTGAGACCTTGAACTTTTAGCTTAAATAATTGCC 960		
QY	961 CAACTTAGTGTCTGGAACACTTACTGACTGTGAATACTGGATCTTGTGAACCTCTTTAA 1020		
Db	961 CAACTTAGTGTCTGGAACACTTACTGACTGTGAATACTGGATCTTGTGAACCTCTTTAA 1020		
QY	1021 GCATCCCTCTGCGCTGGAAGAGTTTGTGGAGGACCTTACACGAGGAACCAAGATA 1080		
Db	1021 GCATCCCTCTGCGCTGGAAGAGTTTGTGGAGGACCTTACACGAGGAACCAAGATA 1080		
QY	1081 CTCCTCTATATCATTTACAGCAAAAGTTATGTCGATGGGTTTAAATATTTGGAAGAA 1140		
Db	1081 CTCCTCTATATCATTTACAGCAAAAGTTATGTCGATGGGTTTAAATATTTGGAAGAA 1140		
QY	1141 TGAGTTGCCCATTTGTTTTCATGTTTGCAGCCGCTTAAATAAATTTGATCTCTATGC 1200		
Db	1141 TGAGTTGCCCATTTGTTTTCATGTTTGCAGCCGCTTAAATAAATTTGATCTCTATGC 1200		
QY	1201 AATGCTAGACACGAGGATCATTTGTTAATCCAAAGGTGCCAAATCTGGAAGTCT 1260		
Db	1201 AATGCTAGACACGAGGATCATTTGTTAATCCAAAGGTGCCAAATCTGGAAGTCT 1260		
QY	1261 TGAGCAAGGAATGTAATTTGGAGATAGAGGTTAGAGTTCTTGGTCGTTGTGAAGAG 1320		
Db	1261 TGAGCAAGGAATGTAATTTGGAGATAGAGGTTAGAGTTCTTGGTCGTTGTGAAGAG 1320		
QY	1321 GCTAAAGGCTTAGATTGAAAGGCGGATCATGATCAAGGAATGGAGGATGAAGAAG 1380		
Db	1321 GCTAAAGGCTTAGATTGAAAGGCGGATCATGATCAAGGAATGGAGGATGAAGAAG 1380		
QY	1381 TACTGTGCTCCATAGAGGCTAATAGCCCTTGTCAAGGCTGTTGAGAGTGAATACAT 1440		
Db	1381 TACTGTGCTCCATAGAGGCTAATAGCCCTTGTCAAGGCTGTTGAGAGTGAATACAT 1440		
QY	1441 GGCTGTTTATGTCGTGATTTACAAATGCTCTTGGACATATTTGGAACTCATTGAA 1500		
Db	1441 GGCTGTTTATGTCGTGATTTACAAATGCTCTTGGACATATTTGGAACTCATTGAA 1500		
QY	1501 GAACTCTCTGATTTTTCGCTTGTGTTGCTTGCATGAAAGAGAGATTAATGATTTGCC 1560		
Db	1501 GAACTCTCTGATTTTTCGCTTGTGTTGCTTGCATGAAAGAGAGATTAATGATTTGCC 1560		
QY	1561 ACTTGCAATAGGGGTGAGGCTCTACTAGGGGCTGTGACAGCTGAGGAGATTGCTCT 1620		
Db	1561 ACTTGCAATAGGGGTGAGGCTCTACTAGGGGCTGTGACAGCTGAGGAGATTGCTCT 1620		
QY	1621 ATATCTCAGGCTGGGGGTTGATGATGAGGCTTGTGTTACATTTGGACAAATACAGTCC 1680		
Db	1621 ATATCTCAGGCTGGGGGTTGATGATGAGGCTTGTGTTACATTTGGACAAATACAGTCC 1680		
QY	1681 AATGTGAGATGGATGCTCTTGGTTTATGTTGGGGAGTCTGATGCGAGGCTTTTGAGTT 1740		
Db	1681 AATGTGAGATGGATGCTCTTGGTTTATGTTGGGGAGTCTGATGCGAGGCTTTTGAGTT 1740		
QY	1741 CGTAAAGGGGTGTCCTAGTCTTCAGAAAATTGAAATGAGAGGCTTTTATTTTTCAGTGA 1800		
Db	1741 CGTAAAGGGGTGTCCTAGTCTTCAGAAAATTGAAATGAGAGGCTTTTATTTTTCAGTGA 1800		
QY	1801 AGTGCACCTGCTGCTGCTGCAACAATTTGATCTCTTAGTCTTGTGGGTGCAAGG 1860		
Db	1801 AGTGCACCTGCTGCTGCTGCAACAATTTGATCTCTTAGTCTTGTGGGTGCAAGG 1860		
QY	1861 TTATGCTGATCTCCATCTGCGCTGATCTTTTGGTAAATGGCTCGACCTTTTGGAAAT 1920		
Db	1861 TTATGCTGATCTCCATCTGCGCTGATCTTTTGGTAAATGGCTCGACCTTTTGGAAAT 1920		
QY	1921 TGAGTTGATTCCTTTAGAAAAGGTGCTACGAATACAATCCAGATGAGACTGATTTGT 1980		
Db	1921 TGAGTTGATTCCTTTAGAAAAGGTGCTACGAATACAATCCAGATGAGACTGATTTGT 1980		







HILAYSLAQRTDCPTTVRVLKEPI"  
1980. .2335

3' UTR  
ORIGIN

Query Match  
Best Local S  
Matches 1196

Query Match 35.2%; Score 805.8; DB 8; Length 2335;  
Best Local Similarity 66.9%; Pred. No. 1.4e-179;  
Matches 1196; Conservative 0; Mismatches 577; Indels 14;

[illegible]

1208	DB	CTTTGCACAGTACTGTAAAGCAGTTGAAGCGGCTTGAGGATTTGAACGGCGTGCAGATGAACA	1267
1360	QY	AGGAATGGAGGATGAAGAAGGTACTGTGTCCCATAGAGGCTAATAAGCCTTGTGCACAGGG	1419
1268	DB	AGGAATGGAGGACGAAGAAGCTTAGTCTCACAAAGAGATTAAATCGCTTTGGCTCAGGG	1327
1420	QY	CTGTTCAGAGCTTTGAATACATACATGGCTGTTTATGTGTCTGATATTACAAATGCATCTCTGGA	1479
1328	DB	CTGCAGGAGCTAGAAATACATGCGCGTGTATGCTCAGATATAACTAACGAATCTCTTTGA	1387
1480	QY	ACATATTGGAACCTCACCTTGAAGAACCTCTGTGATTTTCGCGCTTGTGTCTTGACCATGA	1539
1388	DB	AAGCATAGGCACATATCTGAAAAACCTCTGTGACTTTCCGCGCTTGTCTTACTCGACCGGGA	1447
1540	QY	AGAGAGATAACTGATTTGGCCACTTGACAAATGGGTGAGGGCTCTACTGAGGGGCTGTGA	1599
1448	DB	AGAAAGATTACAGATCTGCCACTTGGACAAACGAGTCCGATCTCTTTTGATTGGATGCA	1507
1600	QY	CAAGCTCAGCAGAGATTTGCTCTATATCTCAGCGCTGGCGGTTGACTGATGATAGGCCCTTGG	1659
1508	DB	GAACTCAGACGATTTGCATTTCTATCTGAGACAAGCGGCTTAAACCGACTTGGCTTAAG	1567
1660	QY	TTACATTTGACAATACACAGTCCAAATGTGAGATGGATGCTGCTGTGGTTATGTGGGGAGTC	1719
1568	DB	CTACATCGGACAGTACAGTCCAAACGTAGATGGATGCTCTGGGTTCACGTAGTGAATC	1627
1720	QY	TGATGCAAGGCTTTTGGAGTTTCCGTAGAGGGTGTCTTAGTCTTTCAGAAAACTTCAAAATGAG	1779
1628	DB	AGATGAAGGTTTAAATGGAAATTTCTCAAGAGGCTGTCCAAATCTACAGAAGCTTAGAGATGAG	1687
1780	QY	AGGGTGTTTATTTTTCAGTCAAAGCTGACATTTGCTGTGGCTTGCAACAAATGACTTCTCT	1839
1688	DB	AGGTTGT--TGCTTTAGTGAGCGAGCAATCGCTGACGGGTTACAAAATTGCCTTCACT	1744
1840	QY	TAGGTACTTTGGGTGCAAGGTTATGGTGTATCTCCA TCTGGACGTGATCTTTTGGTAA	1899
1745	DB	GAGATACCTTTGGGTACAAGTTACAGAGCATCGATGACGGGACAAGATCTAATGCAAT	1804
1900	QY	GGCTCGACCCCTTTTGGAAACATGAGTTGATTCCTTTAGAAAAAGTGGCTAGCAATACCA	1959
1805	DB	GGCTAGACCGTACTTGGAAACATCGAGCTGATTCCAATCAAGAAGATCCCGGAAGTGAATCA	1864
1960	QY	TCCAGATGAGACTGTAGTTGTTTGAGCATCCGCTCATATCTTGCAATATTCTCTTGC	2019
1865	DB	ACAAGSAGAGATAAGAGAGATGGAGCATCCGGCTCATATATGGCTTTACTACTCTTGGC	1924
2020	QY	AGGCGAGAGATCAATTTTCCAGATACTGTTGTGCCCTTTGGACACTGCCACATCGGTGA	2079
1925	DB	TGGCCAGAACAGATTGCTCCAAACA CTGTTAGAGTCTTGAAGAGCCATATGATATG-	1983
2080	QY	TACCTAGAGGCCAGAGCTGTGTATATATATACAGHTTTTCTTTTGT	2126
1984	DB	-ACCAAAAAACAGTTGTATATAAAGATTTTTTAGTCTCGAGTTTT	2029

#### RESULT 4

AV133556	1779 bp	mRNA	linear	PLN 07-AUG-2002
LOCUS	Arabidopsis thaliana At2g39940/728M21.10 mRNA, complete cds.			
DEFINITION	Arabidopsis thaliana At2g39940/728M21.10 mRNA, complete cds.			
ACCESSION	AV133556			
VERSION	AV133556.1 GI:22137081			
KEYWORDS	FUT CDNA.			
SOURCE	Arabidopsis thaliana (thale cress)			
ORGANISM	Arabidopsis thaliana			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
	rosids; eurosids II; Brassicales; Brassicaceae; Arabidopses.			
REFERENCE	1 (bases 1 to 1779)			
AUTHORS	Cheuk, R., Chen, H., Kim, C.J., Shinn, P., Banh, J., Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,			

Onodera,C.S., Palm,C.J., Quach,H.I., Sakurai,T., Satou,M., Seki,M.,  
Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K.,  
Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A.  
and Ecker,J.R.

# Arabidopsis ORF clones

JOURNAL  
REFERENCE  
AUTHORS

Unpublished  
2 (bases 1 to 1779)

Chen,R., Chen,H., Kim,C.J., Shinn,P., Banth,J., Bowser,L.,  
Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,  
Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,  
Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,  
Onodera,C.S., Palm,C.J., Quach,H.I., Sakurai,T., Satou,M., Seki,M.,  
Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K.,  
Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A.  
and Ecker,J.R.

# Direct Submission

JOURNAL

Submitted (17-JUL-2002) Salk Institute Genomic Analysis Laboratory  
(SIGAL), Plant Biology Laboratory, the Salk Institute for  
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,  
USA

# COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the  
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN  
Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J.,  
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,  
Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEN (SSP) Consortium members constructed and  
sequenced the PUNI (ORF) clones using the RAFL cDNAs: Cheuk,R.,  
Chen,H., Kim,C.J., Shinn,P., Banth,J., Bowser,L., Chan,M.M.,  
Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Jones,T.,  
Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Miranda,M., Nguyen,M.,  
Onodera,C.S., Palm,C.J., Quach,H.I., Southwick,A., Tang,C.C.,  
Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Davis,R.W.,  
Theologis,A., and Ecker,J.R.

Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to  
this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)  
contributed equally to this work as PIs.

# FEATURES

source

1..1779  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="2"  
/clone="U12955"  
/ecotype="Columbia"  
/note="This clone is in PUNI 51"

# CDS

1..1779  
/note="LRR-containing F-box protein"  
/codon\_start=1  
/product="At2g39940/T28M21.10"  
/protein\_id="AA091386.1"  
/db\_xref="GI:22137082"  
/translation="MEDPDIKCKLSCVATVDVLEQVMTYLTDPKRDSSASLVCRRW  
FKIDSETHVMAICYPATDRLRRFPNLSLKLKPKRAMNLPENWGGVTP  
WTEISNNLRKQSVHFRMIVSDLDLRLAKARADDLETLLDKSCGFTDGLLSIV  
THCRKTLTLMESSEFSEKDWLHQAQNTSLVFNPMTEFAKIPKDLTARN  
CSRLSVKGVDPLELVEFKAANLEFEFGSLNEDIEMPEKYNLVFPKRLORLG  
LSVMGNEMPIIFPPAAQIRKLDLLVALLTRDHTLIQKCPNLEVLTRNVIGRGL  
EVLQYCKLRLRTERGADEQMEDEGLVSQRLIALAQCELELYMAVVSIDITN  
ESLESIGTYKQCPNRLVLDREKITDPLDNGVRSLLIOCKLRLRPAFYLRQGL  
TDLGSYIKYGVNPMVLLGVYSGESDEGLMEFSRCQPNLQKLEMRGCFSERAAAA  
VTKLPSRLVWQVGRASMTGQDLQMOMAPYNIELIPSRVPEVNNQOEIREMHPA  
HILAYYSLAGQRTDCPTTVRLKEPI"

# ORIGIN

Query Match 35.2%; Score 804.6; DB 8; Length 1779;  
Best Local Similarity 67.8%; Pred. No. 2.6e-179;  
Matches 1160; Conservative 0; Mismatches 539; Indels 12; Gaps 2;  
409 CGTTTCCAGGTGTGTGCGAGCTGGTACGAGCTCGCTACCGCCGCGCTCCCGCAGCGTCAAC 468  
108 GGTCTTTTGTGTGTGGAGATGGTTCAAGATTGATTCGAGACGAGAGCATGTGCAC 167  
469 CATCGCGTCTGTCTACACCAACCCCGGCTCGCTCGCGCGCTCCCGCAGCGTCAAC 528  
168 TATGCGCGTTTGTCTACACTGCGACGCGCTGATCGCTTGTAGCGCTGATTCGCCGAATTGAG 227  
529 GTGCTCAAGCTCAAGGCGAGCCCGGAGCGCGCAATGTTCAACTTGATACCGGAGATTG 588  
228 GTGCTCAAGCTTAAAGCAAGCCCTAGAGCACTATGTTTAACTGATCCCTCGAGAACTG 287  
589 GGGCGGACACGTCACTCCCTCGGTCGCAAGAGATTCTCAGTACTTCGATTCGCTCAAGAG 648  
288 GGGAGGTTATGTTACTCTCTTGGGTTACTGAGATTCTTAACCAACCTTAGCGAGCTCAAAATC 347  
649 CCTCCACTCCGCGCATGATTGTCAGGATTCGAGATTCGAGATTCGCTCGTGACCG 708  
348 GGTGCACTTCCGACGGATGATTGTCAGTACTTAGATCTAGATCGTTTACCTTAAGCTAG 407  
709 CGGTCAAGCTTCAAGCTCTCAAGCTTGACAAGTGTCCGGTTTCAACCGAGTGGTCT 768  
408 AGCAGATGATCTTGAAGCTTGAAGCTAGACAGGTTCTCTGTTTACTACTGATGGACT 467  
769 TTTCATATCGGTGCGCTTTGCAAGAGTTTAAAGTCTTGTGTTTGGAGGAAAGCTCAAT 828  
468 TTTGAGCATCGTTACACACTCGAGGAAATAAATAAATTTGTTAATGGAAGAGATTCTTT 527  
829 TCTTGAGAGACGAGGATCGCTACAGAGCTTGTGTTGTTGAATAATACAGTCTTTGAGAC 888  
528 TAGTGAAAGAGATGTAAGTGGCTTCATGAGCTTCTCAGCAACAACATCTCTTGAGGT 587  
889 TCTCAATTTTACTTTGACAGACATTTGCTTGTGAAGATTGAGGACCTTGAACCTTTTACG 948  
588 TTTAAACTTCTACATGACGAGTTTGCCAAATCAGTCCCAAGACTTGGAAACCATAGC 647  
949 TAAAAATTCGCCAACTTAGTGTCTGTGAAGCTTACTGCTGTAAGTCTGTAATCTGATCTGT 1008  
648 TAGAAATTCGCGCTCTCTGATCTGTGAAGCTCGGTGACTTTGAGATTTTGGAACTAGT 707  
1009 GAACTTCTTTAAGCATGCTCTCGCTCGAAGAGTTTGTGAGGACCACTCAACAGAGA 1068  
708 TGGGTCTTTAAGGCTGACGCTAATCTTGAAGATTTTGTGCTGCTCTTGAATGAGA 767  
1069 -----ACCAGAAAGATCTCTGCTATATCATTAACAGCAAGATTGATCGATGGG 1119  
768 TATTGGAATGCTGAGAAGTACATGAATCTGTGTTTTTCCCGGAAACTATGTGCGCTTGG 827  
1120 TTTAACAATATTGAAAGAAATAGTTCGCCATTTGTGTTTCATGTTGCGCGCTACTAAA 1179  
828 TCTCTCTTACATGGGACCTTAATAATGCCAATCTATTTCCATTCGGCGGCCAAATCCG 887  
1180 AAAATTGGATCTCTCTATGCAATCTAGACGAGGATCATTTGTTATGTTAATCCAAAG 1239  
888 AAGCTGGAATTTGCTTTATGCACTGCTAGAACTCGAAGCACTGACGTTATCCAAA 947  
1240 GTGTCCTAAATCTGGAAGTCTTGTGACAAGGAATGTAATTTGAGATAGAGGTTAGAGT 1299  
948 GTGTCCTAAATTTGGAAGTCTTGTGACAAGGAATGTAATTTGAGATAGAGGTTAGAGT 1007  
1300 TCTTGTGCTGTTTGTGAAGCTTAAAGGCTTAGGATTTGAAGGGCGGATGATGATCA 1359  
1008 CTTTCACAGTACTGTAAAGCTTGAAGCGGCTGAGGATTGAACCGGCTGCAGATGAACA 1067  
1360 AGGAATGGAGGATGAAGAGGTTACTGTGTCCCATAGAGGCTTAATAGCTTGTCAAGG 1419  
1068 AGGAATGGAGGACGAAGAGGCTTAGTCTCAAGAGAGATTAATCGCTTTGCTCGACGG 1127  
1420 CTGTTTCAAGCTTGAATACATGGCTGTTTATGTGCTGATATTAACAATGCACTCTTGA 1479  
1128 CTGCCAGGAGCTAGAAATACATGGCGTGTATGTCTCAGATATAACTAAACGAATCTCTTGA 1187

QY

DB

QY 1480 ACATATTGGAACCTCATTGAAACCTCTGTGATTTTCGCCCTGTGTTGCTTGACCATGA 1539  
 Db 1488 AAGCATAGGCAATATCTGAAAAACCTCTGTGATCTTCGCCCTGTGTTGCTTGACCATGA 1247  
 QY 1540 AGAAGAATAACTGATTTGCCACTTGACAAATGGGGTGAAGGCTCTACTGAGGGGCTGTGA 1599  
 Db 1248 AGAAGGATTAACAGATCTGCCACTGGAACAACGGAGTCCGATCTCTTTGATTGGATGCA 1307  
 QY 1600 CAAGCTGAGGAGATTGCTCTATATCTCAGGCGTGGCGGGTGTGATGTAGGCTTGG 1659  
 Db 1308 GAAACTCAGACGATTGCACTTATCTAGAGCAAGCGCGCTTAACCGACTTGGGCTTAAG 1367  
 QY 1660 TTACATTGCAATACAGTCCAAATGTGAGATGGATGCTGTGTTATGTGGGGAGTC 1719  
 Db 1368 CTACATCGACATGACATCTCAAGCTGAGATGGATGCTGTGGTACGTAGGTGATC 1427  
 QY 1720 TGATGACGGCTTTTGGAGTTGCTTAAGGGTGTCTCTAGTCTTTCAGAACTTGAATGAG 1779  
 Db 1428 AGATGAAGGTTTAAATGGAATCTCAAGAGGCTGTCCAAATCTACAGAACTAGATGAG 1487  
 QY 1780 AGGGTGTATTTTTCAGTGAAGCTGCACTGTGCTGTGGCTGCAACACAACTGACTTCTCT 1839  
 Db 1488 AGGTTGT---TGCTTCAGTGAAGCAATCTGCTGAGCGGTTACAAAATTGCTTCTACT 1544  
 QY 1840 TAGGTACTTGTGGTGCAAGGTTATGTTGTTATCTCCATCTGACGCTGATCTTTTGTGTAAT 1899  
 Db 1545 GAGATACTTGTGGGTACAAAGTTACAGATCTGATGAGCGGACAAAGATCTAATGCAGAT 1604  
 QY 1900 GGCTGACGCTTTTGGAAATTCAGTTGATTCCTTCTAGAAAGTGGCTACGAATACCAA 1959  
 Db 1605 GCTAGACGCTTGTGAATCATGAGCTGATTCATCAAGAGAGTCCCGGAAGTGAATCA 1664  
 QY 1960 TCCAGATGAGCTGTGTTGTTGAGATCTCTCTATATCTTGTATATATCTTCTTCTTGC 2019  
 Db 1665 ACAAGGAGATGAAGAGATGAGCATCCGCTCATATATTTGGCTTACTACTCTCTGTC 1724  
 QY 2020 AGGCGAGATCAGATTTTCCAGATCTGTT 2050  
 Db 1725 TGGCCAGAGAAACAGATTGTCCAACTGTT 1755

RESULT 5  
 AY168645  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 1 (bases 1 to 2088)  
 Wang, W. and Wang, X.  
 Direct Submission  
 Submitted (25-OCT-2002) Biochemistry, Fudan University, 220 Handan  
 Road, Shanghai 200433, P.R. China  
 Location/Qualifiers  
 1. .2088  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="japonica"  
 /db\_xref="taxon:39947"  
 /chromosome="1"  
 66. .1853  
 /note="rCOI1; similar to Arabidopsis thaliana COI1"  
 /codon\_start=1  
 /product="COI1"  
 /protein\_id="AA038719.1"  
 /db\_xref="GI:37359393"  
 /translations="MGSEVPEPRRLNRALSFDWVPEALHLVGHVDEPRDREASR  
 VCRWRHIDALTRKHVTVAFCYAARPARLRFRPRLSLKPKPRAAYGLIPDDWG

## FEATURES

source  
 1. .2088  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="japonica"  
 /db\_xref="taxon:39947"  
 /chromosome="1"  
 66. .1853  
 /note="rCOI1; similar to Arabidopsis thaliana COI1"  
 /codon\_start=1  
 /product="COI1"  
 /protein\_id="AA038719.1"  
 /db\_xref="GI:37359393"  
 /translations="MGSEVPEPRRLNRALSFDWVPEALHLVGHVDEPRDREASR  
 VCRWRHIDALTRKHVTVAFCYAARPARLRFRPRLSLKPKPRAAYGLIPDDWG

## CDS

QY 1179 AAAAAATTGATCTCTCTATGCAATGCTAGACACGAGATCATTTGTATGTTAATCAAA 1238  
 Db 1179 AAAAAATTGATCTCTCTATGCAATGCTAGACACGAGATCATTTGTATGTTAATCAAA 1238  
 QY 1119 GTTTAAACATATATGGAAGAATGAGTTGCCCAATGTGTTCATGTTTGCAGCGGTCTAA 1178  
 Db 901 GGTCTACCTACATGGGAACAATGAGATGCTGTATCTTCCCTTTTTCGATGAAACTCA 960  
 QY 841 AGGTAGGAGCTCACCAAGTATGAAAAGTTAAGTTCCACCCAGATATGCTTCTTGG 900  
 Db 1065 AGGAACCCGA-----AAGATACTCTGTATATATACCAAGAAAGTTATGTCGATTGG 1118  
 QY 781 TGATTAGTTTTTTTCAACACAGCAATGCGCTGCAAGACTTTCGTGAGAGAGCATTCACG 840  
 Db 1005 TTGTGAATCTTTTAAAGCATGCTCTGCGCTGGAAGAGTTTGTGAGGACCACTCAACG 1064  
 QY 721 TTGCAAGAATTTGCAAGTCAATTTGATTTCAATGAAGATGATGATGATGATCTTCAGATC 780  
 Db 664 TGACACTGAATTTTACATGACTGAATCT---CAAGTGGCGGCAGCTGATCTAGACTTC 720  
 QY 945 TAGCTAAAATTTGCCCAACTTAGTGTCTGTGAAACTTTACTGACTGTGAAAATCTGGATC 1004  
 Db 825 CAATTTTGAAGACCGAGAAATGGCTACAGAGCTGCTTGTGAATAATACAGTCTTGG 884  
 QY 604 ATATTACTGATAAGGGTGGTGAATGGCTTCATGAACCTTGTCTGTCAACAATCTGTCTGG 663  
 Db 885 AGACTCTCAATTTTACTTACAGACATGCTGTGTGTAAGATTTGAGGACCTTGAACTTT 944  
 QY 664 TGACACTGAATTTTACATGACTGAATCT---CAAGTGGCGGCAGCTGATCTAGACTTC 720  
 QY 945 TAGCTAAAATTTGCCCAACTTAGTGTCTGTGAAACTTTACTGACTGTGAAAATCTGGATC 1004  
 Db 721 TTGCAAGAATTTGCAAGTCAATTTGATTTCAATGAAGATGATGATGATGATCTTCAGATC 780  
 QY 1005 TTGTGAATCTTTTAAAGCATGCTCTGCGCTGGAAGAGTTTGTGAGGACCACTCAACG 1064  
 Db 781 TGATTAGTTTTTTTCAACACAGCAATGCGCTGCAAGACTTTCGTGAGAGAGCATTCACG 840  
 QY 1065 AGGAACCCGA-----AAGATACTCTGTATATATACCAAGAAAGTTATGTCGATTGG 1118  
 Db 841 AGGTAGGAGCTCACCAAGTATGAAAAGTTAAGTTCCACCCAGATATGCTTCTTGG 900  
 QY 1119 GTTTAAACATATATGGAAGAATGAGTTGCCCAATGTGTTCATGTTTGCAGCGGTCTAA 1178  
 Db 901 GGTCTACCTACATGGGAACAATGAGATGCTGTATCTTCCCTTTTTCGATGAAACTCA 960  
 QY 1179 AAAAAATTGATCTCTCTATGCAATGCTAGACACGAGATCATTTGTATGTTAATCAAA 1238

## ORIGIN

Query Match 30.0%; Score 685.8; DB 8; Length 2088;  
 Best Local Similarity 63.7%; Pred. No. 3.6e-151;  
 Matches 1111; Conservative 0; Mismatches 617; Indels 15; Gaps 4;

QY 345 GTGTGGTGCAGTGTCTCGACTGCGTCACTCCCTTATCAGACAGACCCCAAGGACCGCG 404  
 Db 124 GGGTCCCCGAGCGCGCTGCACCTCTGTGATGGGCCAGCTCGAGGACCGCGGGACAGG 183  
 QY 405 ACGCCGTTTCCACAGTGTGTGACGCTGGTACGAGCTCGACTCGCTACCCGCAAGCAG 464  
 Db 184 AGCGGCGCTCGCGGGTGTGCGCGCTGCGACCGCATCGACGCGCTCACGCGCAAGCAG 243  
 QY 465 TCACATCGCGCTCTGTCTACACACACCCCGGCTCGCTCGCGCGCGCTTCCCGCACC 524  
 Db 244 TCACGCTCGCTTCTGTCTACGCGCGCGCGCGCGCTTCCGGAGCGGTTCCCGCGGC 303  
 QY 525 TCGAGTCCCTCAAGCTCAAGGCGAGCGCGCGCAATGTTCAACTTGATACCGGAG 584  
 Db 304 TCGAGTCCCTCTCGCTCAAGGCGAGCGCGCGCGCATGTACGGGCTCATCCCGGAG 363  
 QY 585 ATTGGGGCGGACACATCTCACTCCCTGGGTCAAGAGATTTCTCAGTACTTTCGATTGCTCA 644  
 Db 364 ACTGGGGCGCTACCGCGCGCATGATCGACGAGCTCGCGCGCGCTCGAGTGCCTCA 423  
 QY 645 AGAGCTTCCATTTCCGCGCATGATGTCAAGATTCGATCTTCAGATCTTCGCTCGTG 704  
 Db 424 AGGCGCTCCACCTCGCGCGCATGACCGTCAACGACCGCGCATCGCGCGCTTGTCCGCG 483  
 QY 705 ACCGCGGTCAAGTCTTCAAGCTTCAAGTGTGCAAGTGTCCGTTTCAACCGGATG 764  
 Db 484 CCGCGGACACATGCTGCGAGGAGCTCAAGCTCGAAGTGCATGCGCTTCTCCACTGACG 543  
 QY 765 GTCTTTTCCATATCCGTCGCTTTTCAAGAGTTTAAGAGTCTTGTTTTGGAGGAAAGT 824  
 Db 544 CCCTCGGCTCTCGCGCGCTCGTCAGATCCCTGAGAACTTTATTTCTGGAAGAGTGC 603  
 QY 825 CAATTTTGAAGACCGAGAAATGGCTACAGAGCTGCTTGTGAATAATACAGTCTTGG 884  
 Db 604 ATATTACTGATAAGGGTGGTGAATGGCTTCATGAACCTTGTCTGTCAACAATCTGTCTGG 663  
 QY 885 AGACTCTCAATTTTACTTACAGACATGCTGTGTGTAAGATTTGAGGACCTTGAACTTT 944  
 Db 664 TGACACTGAATTTTACATGACTGAATCT---CAAGTGGCGGCAGCTGATCTAGACTTC 720  
 QY 945 TAGCTAAAATTTGCCCAACTTAGTGTCTGTGAAACTTTACTGACTGTGAAAATCTGGATC 1004  
 Db 721 TTGCAAGAATTTGCAAGTCAATTTGATTTCAATGAAGATGATGATGATGATCTTCAGATC 780  
 QY 1005 TTGTGAATCTTTTAAAGCATGCTCTGCGCTGGAAGAGTTTGTGAGGACCACTCAACG 1064  
 Db 781 TGATTAGTTTTTTTCAACACAGCAATGCGCTGCAAGACTTTCGTGAGAGAGCATTCACG 840  
 QY 1065 AGGAACCCGA-----AAGATACTCTGTATATATACCAAGAAAGTTATGTCGATTGG 1118  
 Db 841 AGGTAGGAGCTCACCAAGTATGAAAAGTTAAGTTCCACCCAGATATGCTTCTTGG 900  
 QY 1119 GTTTAAACATATATGGAAGAATGAGTTGCCCAATGTGTTCATGTTTGCAGCGGTCTAA 1178  
 Db 901 GGTCTACCTACATGGGAACAATGAGATGCTGTATCTTCCCTTTTTCGATGAAACTCA 960  
 QY 1179 AAAAAATTGATCTCTCTATGCAATGCTAGACACGAGATCATTTGTATGTTAATCAAA 1238

961	AGAAACTGGACTTGGCAATACACTTTTCTCACAACAGAGATCATTTGTACAGATTATTCGAA	1020
1239	GGTGTCCAAATCTGGAAAGTCCCTTGACACAGGAATGTAATTCGAGATAGAGGTTAGAGG	1298
1021	AATGTCCCCTATCTAATTTCTTGAGTGAAGAACGTGATAGAGATAGAGGCTAGAAG	1080
1299	TTCTTGGTTCGTTTGTGAAGAGGCTAAAAGGCTTAGGAATTGAAAGGGCGATGATGATC	1358
1081	TTCTTGGTGATACATGCAAGAAGCTACGAAGCTCCGAATTCGACGGGCTGATGATGATC	1140
1359	AAGGAATGGAGGATGAAGAAGGTTACTGTGTCCATAGAGGGCTAATAGCTTTGTACAGG	1418
1141	CAGGTCTGCAGGAAGACGACGAGGAGGATTTCTCAGCTAGGCTTGACAGCGCTGTCTTGG	1200
1419	GCCTGTTCCAGAGCTTGAATACATCGCTGTTTATGTGTCTGATATTACAAATGCATCTCTGG	1478
1201	GTTCGCGTGAATTGGAGTACATAGCTGCCATGCTATCGGATATCACCAATGGGCGCTTGG	1260
1479	AACATATTGGAACCTCACTTGAAGAACCTCTGTGATTTTCGCCCTGTGTTGCTTTGACCATG	1538
1261	AGTCTATTGGACTTTCTGCAAAATCTATAOCAGCTTTTGGCTTGTGCTACTTGAACAGAG	1320
1539	AAGAGAAGATAACTGATTGGCCACTTGCAAAATGGGGTGAGGCTCTACTGAGGGGCTGTG	1598
1321	AAAGACAGGTAAACAGATCTGCCATTGCACAAATGGTGTCTGTGCTGTTTAAGAAAATTGCA	1380
1599	ACAGCTGAGGAGATTGGCTCTATATCTCAGGGGTGGCGGTTGACTGATGTAGGCGCTTG	1658
1381	CAAAGCTTCGGAGGTTTGCTCTTACTCTTACACAGGAGGCTTTGATGATGGGCTTGA	1440
1659	GTATACATTGGACAATACAGTCCAAATGTGAGATGGATGCTGCTTGGTTATGTGGGGAGT	1718
1441	GCTATATCGGACAGTACAGTGGAAATATCCAATACATGCTACTGGGCAATGTTGGTGAAT	1500
1719	CTGATGCGAGGCTTTTGGAGTTCGCTAAGGGGTGCTCTAGTCTTTCAGAACTTTGAAATGA	1778
1501	CTGACCATGGATTGATCCGTTTTCGACGTGGGCTGCACCAACCTTTCAGAACTTTGAAATGA	1560
1779	GAGGGTGTATTTTTCAGTGAACTGCTACTGTGTGGCTGCAACACAAATTCGACTTCTC	1838
1561	GAAGCTG--CTGGCTTCAGCGAGCGAGCTTTGTCCCTGCTGTACTCGAGTGCCTTCCC	1617
1839	TTAGTACTTCTGGGTCAAGGTTATGGTGTATCTCCATCTGGACGCTGATCTTTTGGTAA	1898
1618	TGCAATACATATGGTGCAGGATACAGAGCATCTCAAACAGGCGCTTGACCTCTGCTCA	1677
1899	TGGCTCGACCCCTTTTGGAACTGAGTTGATTCCTCTTAGAAAGGTGGCTACGAAATACCA	1958
1678	TGGCCAGGCGCTTCTGGAACATCAGATTTTACACCTCCGAGCGCTTGAGAGTTTAAATCAT	1737
1959	ATCCAGATGAGACTGTGA--GTTCTGTGAGCATCTGCTCATATCTTTGCATATTATTCTC	2015
1738	TGACAGAAAGATGGAGAACCCCTGTGTGGATAGCCATGCTCAGGTTTCTTGCCTACTATTCCC	1797
2016	TTGAGGGCGACAGATCAGATTTTCCAGATCTGTTGTGCTTTTGGACACTGCCCATGCG	2075
1798	TTGCTGGGAAGAGGCTCTGCTGCCCTCAGTGGGTGATCCCTTGTGATCTCCGTGATTGT	1857
2076	TTG 2078	
1858	TTG 1860	

RESULT 6	AF536527	711 bp	DNA linear	PLN 01-OCT-2002
LOCUS	AF536527			
DEFINITION	AF536527			
	AF536527.1			
ACCESSION	AF536527			
VERSION	AF536527.1			
KEYWORDS				
SOURCE				
ORGANISM				

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE  
1 (bases 1 to 711)  
Christiansen,L.C. and Ulvskov,P.  
Putative Glycine max sbCOI1 homologous to Arabidopsis COI (atFBL2)  
Unpublished  
2 (bases 1 to 711)  
Christiansen,L.C. and Ulvskov,P.  
Direct Submission  
Submitted (08-AUG-2002) Biotechnology Group, Danish Institute of  
Agricultural Sciences, Thorvaldsensvej 40, Frederiksberg C 1871,  
Denmark

FEATURES  
source  
1..711  
/organism="Glycine max"  
/mol\_type="genomic DNA"  
/cultivar="FGX1835-2E"  
/db\_xref="taxon:3847"  
!..>711  
/gene="COI1"  
/note="sbCOI1"  
!..>711  
/gene="COI1"  
/products="putative coronatine-insensitive 1"  
!..>711  
/gene="COI1"  
/note="sbCOI1; similar to Arabidopsis COI1 (atFBL2) "  
/codon\_start=1  
/product="putative coronatine-insensitive 1"  
/protein\_id="AA031713.1"  
/db\_xref="GI:23395153"  
/translation="SSMQAVSHRGTLALSGCGSELYMAYVYDINASLEHIGTHLK  
NLCDRLVLDHEEKIDLPDNGVALLRGDCKLRPALYLRGGITDVGIVGYQY  
SNVRWMLGYVGEISDAGLEFAKGCPSKUEMRGCLFFSERALAVAAATQLTSRLY  
WVQGYGSPSGDRLLVWVPMFNIELIPSRKVATNTNPDETVVVEHPAHILAYISLGL  
ORSDPDTVVVLDATWMLD"

gene  
mRNA  
CDS

## ORIGIN

Query Match	29.9%	Score	583.2	DB	8	Length	711
Best Local Similarity	99.6%	Pred. No.	1.2e-150				
Matches	685	Conservative	0	Mismatches	3	Indels	0
Gaps	0						
QY	1383	CTGTGCCCCATAGAGGGCTAATAGCCTTGTCACAGGGCTGTCAGAGCTTGAATACATGG	1442				
Db	14	CGGTGCCCATAGAGGGCTAATAGCCTTGTCACAGGGCTGTCAGAGCTTGAATACATGG	73				
QY	1443	CTGTTTATGTGCTGATATTACAAATGCATCTCTGGACATATTGGAACATCAGTTGAAGA	1502				
Db	74	CTGTTTATGTGCTGATATTACAAATGCATCTCTGGACATATTGGAACATCAGTTGAAGA	133				
QY	1503	ACCTCTGTGATTTTGGCCTTGCTGTGTCACCAATGAAGAGAAGATAACTGATTTGCCAC	1562				
Db	134	ACCTCTGTGATTTTGGCCTTGCTGTGTCACCAATGAAGAGAAGATAACTGATTTGCCAC	193				
QY	1563	TTGCAATAGGGGTGAGGGCTCTACTAGGGGCTGTGCAAGCTTGAGAGATTTGCTCTAT	1622				
Db	194	TTGCAATAGGGGTGAGGGCTCTACTAGGGGCTGTGCAAGCTTGAGAGATTTGCTCTAT	253				
QY	1623	ATCTCAGGCTGGGGGTTCAGCTGATGTAGGCCCTTGTTTACATTGGACAATACAGTCCAA	1682				
Db	254	ATCTCAGGCTGGGGGTTCAGCTGATGTAGGCCCTTGTTTACATTGGACAATACAGTCCAA	313				
QY	1683	ATGTGAGATGGATGCTGCTCTTGTTATGTGGGGAGTCTGATGCAAGGCTTTTGGAGTTG	1742				
Db	314	ATGTGAGATGGATGCTGCTCTTGTTATGTGGGGAGTCTGATGCAAGGCTTTTGGAGTTG	373				
QY	1743	CTAAGGAGTGCTCTAGTCTTCAGAACTTCGAAATGAGAGGGTGTATTATTTTTCAGTGAAC	1802				
Db	374	CTAAGGAGTGCTCTAGTCTTCAGAACTTCGAAATGAGAGGGTGTATTATTTTCAGTGAAC	433				
QY	1803	GTGCACCTTGCTGGGCTGCACACAAATTGACTCTCTTATGAGTACTGTGGGTGCAAGTT	1862				



QY	585	ATTGGGCGGACACGTCACCTCCCTGGGTCAAGAGATTTCTCAGTACTTCGATTCGCTCA	644
Db	380	ACTGGGCGGCTACGCGCGCCATGGATCGACAGCTCGCGCGCGCTCGATGGCTCA	439
QY	645	AGAGCTCCACTTCGCGCGCATGTTCTCAAGATTTCCGATCTTTCAGATCTCGCTCGTG	704
Db	440	AGGCGTCCACTTCGCGCGCATGACCGGTACCGACCGCGCATCGCGCTTCGCGCG	499
QY	705	ACCGCGTCACTGCTTCACGCTCTCAAGCTTGACAGAGTCTCGGTTTACACCGATG	764
Db	500	CCCGCGACATCTCGCAGGCTCAAGCTGACAGAGTCTCGGCTTCACACGATG	559
QY	765	GTCTTTTCATATCGGTGCTTTTGAAGAGTTTAAAGTCTTTTTCGAGGAAAGCT	824
Db	560	CCCTCCGCTCGTCCGCGCTCGTGCAGATCCCTTGAGACTTTATTTCTGGAAGAGTGC	619
QY	825	CAATCTTGAGAGGACGGAATGGGTACACGAGCTTGCTTTGAAATTAATACAGTCTTG	884
Db	620	ATAATTAAGTAAAGGTGGTGAATGCTTCATGAATCTGCTCAAAATTTCTGCTTG	679
QY	885	AGACTCTCAATTTTACCTGACAGACATGCTGTTGTAAGATTTGAGAGCTTCAACTTT	944
Db	680	TGACACTGAATCTTACATGACTGACT---CAAGTGGCGCGAGCTGATCTAGAGCTTC	736
QY	945	TAGCTAAATAATGGCCCACTTAGTGTCTGAAACTTACTGACTGTGAAATACTGGATC	1004
Db	737	TTGCAAGAAATTCGAAGTCAATGATTTCAATGAAATGAGTGTGATCTTTTCAGATC	796
QY	1005	TTGTGAATCTTTTAAGCATGCTCTGCGCTCGAAGATTTTGTGAGGACCTTCAACG	1064
Db	797	TGATTAATGTTTTTCAACAGCAATGCGCTGCAAGACTTTGCTGAGAGCATTTCTACG	856
QY	1065	AGGAACCAAG-----AAGATCTCTGCTATATCATTTACACGAAAGTTATGCTGATGG	1118
Db	857	AGGTAGGAGCTCACCAAGTATGAAAGTTAAGTTTCCACCCGAGTTATGCTTCTGG	916
QY	1119	GTTTAATATATTTGAAGAAATGAGTGGCCATTTGTTGATTTGTTGAGCGGCTACTAA	1178
Db	917	GGCTTACCTACATGGAACAAATGAGATGCTTTATCTCCCTTTTCGATGAAACTCA	976
QY	1179	AAAAATGGATCTCTATGCAATGCTAGACAGGAGATTTGTTGTTAATCCAA	1238
Db	977	AGAACTGGACTTGCATACATTTTCTCAACAGAGATCATTTGTGAGATTTGCA	1036
QY	1239	GGTGTCAAAATCTGGAAGTCTTGAGACAGGATGTAATGAGATAGAGGTTAGAGG	1298
Db	1037	AATGTCCAAATCTACTAATTTCTTGAGGTGAGAACGTTGATAGGATAGAGGCTAGAAG	1096
QY	1299	TTCTTGTGCTTTGTTGTAAGAGCTTAAAGAGCTTAGGATGTAAGAGGCGATGATC	1358
Db	1097	TGTTGTGATGATGATGAGAGCTACGAAGCTCCGAATTCGAATTCGATGATC	1156
QY	1359	AAGGAATGGAGATGAAGAAGTACTGTGCTCCATAGAGGCTAATAGCTTTGTCACAGG	1418
Db	1157	CAGTCTGACGAGGACGAGGAGTTCTCAGTAGGCTTGACAGCCGTTGCTGTTG	1216
QY	1419	GCTGTTAGAGCTTGAATACATGCTGTTTATGTTGCTGATATTAATAATCATCTCTGG	1478
Db	1217	GTTGCCGTGAATTTGAGTACATAGCTGCTGCTATGATCGATATCACCAATGGGCGCTGG	1276
QY	1479	ACATATTCGAATCTCACTTGAAGAACCCTCTGTGATTTTGGCTGTTGTTGTTGACCATG	1538
Db	1277	AGTCTATGGGACTTTCTGCAAAATCTATACGACTTTTCGGCTTGTGCTTGTGACAG	1336
QY	1539	AAGAGAGATTAATGATTTGCACTTGCAATGGGTTAGGCTCTACTGAGGCGCTGTG	1598
Db	1337	AAAGACAGGTAAACAGATCTGCCACTTGACAATGGTGTCTGCTCTGTTAAGAAATGCA	1396
QY	1599	CAAGCTGAGGATTTGCTCTATATCTCAGGCGGTGGCGGTTGACTGATGTAGGCTTTG	1658
Db	1397	CAAGCTTCGAGGTTTGTCTCTATACCTTAGACGAGGAGGCTTTTCAGATGATGGCTTA	1456

593 GGACAGCTACTCCCTGGGCTCAAGAGATTTCTCAGTACTTTCGATTGCGCTCAAGAGCCTC 652  
 530 GCTACGCGCGCCCTGGGTCGCGAGCTCGCGCGCGCTCGAGTCCCTCAAGCGCTC 589  
 653 CACTTCGCGCGCATGATTTGTCAGAGATTCGATCTTCAGAAATCTGCTCGTGAACCGGGT 712  
 590 CACTCGCGCGCATGATTTGTCAGAGATTCGATCTTCAGAAATCTGCTCGTGAACCGGGT 649  
 713 CAGTGTCTTCAAGCTCTCAGAGCTTGAAGTGTCTCGGTTTTCACCCAGGATGCTTTTTC 772  
 650 CACATGCTGAGAGCTCAAGCTCGAAGTGTCTCGGTTTTCACCCAGGATGCTTTTTC 709  
 773 CATATCGGTCGCTTTTTCGAAGATTTAAGAGTCTTTGTTTGGAGGAAAGCTCAATTTCT 832  
 710 CTGCTGCGCGCTCTCGATCACTCAGAACATTTATTTCTGGAGGAATGCTCAATTTCT 769  
 833 GAGAGGACGAGAAATGGCTACAGAGCTTGTGTTTGAATATATACAGTCTTGGAGACTCTC 892  
 770 GATAATGGTACTGAATGGCTCCACGAGCTTGTCTCAACAATCTGTTCTGGAGACATTG 829  
 893 AATTTTCTTACAGACATGCTGTTGTGAAGATTGAGGACCTTGAACCTTTAGCTAAA 952  
 830 AACTTCACATGACCGAACTCACAGTGGT---GCCAGCTGACCTGGAGCTTCTCGAAG 886  
 953 AATTCGCCCAACTTAGTGTCTGTGAAGTCTTACTGACTGTGAAATPACTGGATCTTGTGAAC 1012  
 887 AAGTGAAGTCACTAATTTCTTGAAGTCACTGACTGTGACTTTTTCAGATTTAATTTGA 946  
 1013 TTTCTTAAAGATGCTCTGCGCTGGAGAGATTTTGTGGAGGACCTTACAGGAGAACCA 1072  
 947 TTTTTCGGATGCTGATCATTTTCAAGAGATTTTGGGAGGGGATTCATTTGAGCAAGG 1006  
 1073 GA-----AGATACTCTGCTATATCATTTACACAGAAAGTTATGTCGATGGGTTTAA 1126  
 1007 GAGCTCACTAAGTATGAATGTTTAAATTTCCCTTCAAGACTGTGCTCTCTTAGGACTTAG 1066  
 1127 TATATTGGAAGATGATGCTCCCATTTGTTTCTATGTTTTCAGCGCTGACTTAAATAAATG 1186  
 1067 TACATGGGCAACAGAGATGCCATTTATTTCTCTTCTCTGATTTACTTCAAGAGCTG 1126  
 1187 GATCTCTCTAAGCAAGCTAGACAGGAGATCATTTGTATGTTTAAATTCAGAGTGTCCA 1246  
 1127 GACTTGCAGTACACTTTTCTCACCCTGAAAGTCACTGCCAACTCATTTGCAAAATGCTCC 1186  
 1247 AATCTGGAAGTCTTTGAGCAAGGAATGATTTGAGATAGAGGCTTAGAGGTTCTTGGT 1306  
 1187 AACTTACTAGTTCTTTCGGTGAAGATGATTTGAGATAGAGATAGAGGTTGTTGCA 1246  
 1307 CGTTGTTTGAAGGCTTAAAGGCTTAGGATTTGAAAGGGGCGATGATGATCAAGGAATG 1366  
 1247 GACACATGCAAGAGCTTCAAGAGCTCAGAGTTGAGCGAGGAGATGATGATCCAGGTTTG 1306  
 1367 GAGATGAGAGAGTACTGTTGCTCCATAGAGGCTTAAAGCTTGTCTACAGGCTGTTTCA 1426  
 1307 CAAGAAGAACAGGAGGAGTCTCTCAAGTCGGTTGTGCAACTGTGATGCGGTAGGATGCGCT 1366  
 1427 GAGCTTGAATACATGCTGTTTATGTTGCTGATATTAACAATGATCTCTGGAACATATT 1486  
 1367 GAATGGGAATACATAGCTGCTTATGTTGCTGATATTAACAATGATGCGGCTGAGTCTATT 1426  
 1487 GGAACTCACTTGAAGAACTCTGTGATTTTCCCTTGTGCTTGTGCTTGCATGAGAGAGG 1546  
 1427 GGAATCTTCTGCAAAATCTTTTGCACCTTCCGCTTGTGCTTGCATGAGAGAGAGG 1486  
 1547 ATAACTGATTTGCCATCTGACAAATGGGTGAGGCTCTACTGAGGCTGTGACAGCTG 1606  
 1487 ATAACTGATTTGCCATCTGACAAATGGGTGAGGCTCTACTGAGGCTGTGACAGAACTT 1546  
 1607 AGGAGATTTGCTTATCTTACAGGCTGCGGCTTGTGATGATGATGAGGCTTGGTTACATT 1666  
 1547 CGAGGATTTGCTTATCTTACAGGCTGCGGCTTGTGATGATGATGAGGCTTGGTTACATT 1606  
 1667 GGACAATACAGTCCAAATGTGAGATGATGCTGCTTGGTTTATGTTGGGGGAGTCTGATGCA 1726

1607 GGACAGTACAGTGAATTTATCAATACATGCTTCTGGTAAATGTTGGGGAACAGATGAT 1666  
 1727 GGGCTTTTGGAGTTCCTAAGGGTGTCTAGTCTTTCAGAAAATTTGAAATGAGAGGTTGT 1786  
 1667 GGTCTGATCGGTTTCGATTTGGGTTGTGAGAACTTCGGGAAGCTTGA---GCTAAGGAGT 1723  
 1787 TTAATTTTTCAGTGAACGTCGACTTGTGTGGCTGCAACACAAATGATTTCTCTTAGTAC 1846  
 1724 TGTGTCTTCACTGAGCAAGCTTTAGCCCGCTTATACGGAGTATGCTTCCCTGAGATAC 1783  
 1847 TTGTGGGTCCAAAGTATGTTGTTATCTCCATCTGGAGCTGATCTTTTGGTAAATGGCTCGA 1906  
 1784 GTGTGGGTACAGGCTTACAGGCTTTTAAGACTGCTTCAAGATCTCATGCTCATGCGCCAGG 1843  
 1907 CCCTTTTGGAACTTCAAGTGTGATTTCTTTCAGAAAAGGTGGCTTACGAATACCAATCCAGAT 1966  
 1844 CCCTTCTGNAACATAGATTTTACCTCCAGTCTGAGAAATCGAAATCGAATGAGAGAA 1903  
 1967 GAGACTGTA---GTTGTTGAGCATCTGCTCATATTTCTTGCATATTTATCTCTTGCAGGG 2023  
 1904 GATGGTGAACCTTGTGTAGATAGTCAAGCTCAGATCTTGCATATCTACTCCCTTGCCTGG 1963  
 2024 CAGAGATCAATTTTCCAGATACTGTTGTGCTTGGACTTGGACACTGC 2067  
 1964 AAGAGGTCGAGCTGCCACAGATCTGTGTTTCTTGTATCTCTGC 2007

## RESULT 9

## AKI01514

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

2305 bp mRNA linear PLN 24-JUL-2003  
 Oryza sativa (japonica cultivar-group) cDNA clone:J033046F05, full  
 insert sequence.  
 AKI01514  
 AKI01514.1 GI:32986723  
 FULL-CDNA; CAP trapper.  
 Oryza sativa (japonica cultivar-group)  
 Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzeae; Oryza.  
 1  
 The Rice Full-length cDNA Consortium, National Institute of  
 Agrobiological Sciences Rice Full-length cDNA Project Team.  
 Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,  
 Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,  
 Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,  
 Ohtsuki, K., Shishiki, T., Foundation of Advancement of International  
 Science Genome Sequencing & Analysis Group, Ootomo, Y., Murakami, K.,  
 Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Taunoda, Y.,  
 Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,  
 Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,  
 Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,  
 Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN,  
 Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,  
 Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,  
 Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Ootani, N., Ota, Y.,  
 Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
 Yoshino, M., and Hayashizaki, Y.  
 Collection, mapping, and annotation of over 28,000 cDNA clones from  
 japonica rice  
 Science 301 (5631), 376-379 (2003)  
 22752273  
 12869764  
 2 (bases 1 to 2305)  
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,  
 Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,  
 Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,  
 Hori, F., Hotta, I., Iida, J., Ikeda, R., Imamura, K., Imoto, A., Itoh,  
 Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,  
 Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,  
 Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,  
 Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,



Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Nariakawa, R., Nikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otomo, Y., Ry, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shingawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, P., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Teunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., and Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission  
Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica rice.  
URL: <http://cdna01.dna.affrc.go.jp/cDNA/>  
NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

PAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariakawa, R., Nikura, J., Oka, M., Ryu, K., Sugano, S., Sugiyama, A., Suzuki, Y., Teunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Oota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, P., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayaahizaki, Y.

Location/Qualifiers  
1. 2305  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="J033046F05"

FEATURES  
source  
Query Match 29.1%; Score 666.8; DB 8; Length 2305;  
Best Local Similarity 63.3%; Pred. No. 1.1e-146;  
Matches 1092; Conservative 0; Mismatches 617; Indels 15; Gaps 4;  
QY 353 GACGTGTCCTGACTGGTGCATCCCTTACATCGACGACCCCAAGGACCGGACCGGTT 412  
DB 287 GAGGAGGCGCTGCACCTGGTCTGGGTACGTGGACGACCCCGGAGCAGGAGCGGTG 346  
QY 413 TCCAGAGGTGTGTGACGCTGTGTGAGCTGCTGCTCACCCTGACGACGCTCACCATC 472  
DB 347 TCGTCTGTGTGCGCGCGCTGCGACCGCATCGACGGCTCAGCGGAGCAGCTCACCCTG 406  
QY 473 GCGCTCTGCTACACCAACCCCGGCTCGCTTCGCGCGCGCTTCGCGCACTCGAGTCG 532  
DB 407 CCTTCTGCTACGCGCGCTGCGCGCGCACCTGCTCGCGCGCTTCGCGCGCTGAGTCG 466  
QY 533 CTCAGCTCAAGGGCAGCGCGCGCGCAATGTTCACTTGATACCCGAGGATTGGGCG 592

467 CTCGCGGTCAAGGGGAAGCGCGCCGCCATGTACGGGCTCATCCCGAGAGACTGGGCG 526  
QY GGACACGTCACCTCCCTGGGTCAAGAGAGATTTCTCAGTACTTTCGATTTGCTCAAGAGCCTTC 652  
DB GCCTACGGCGCGCTGGTTCGCGAGCTCGCGCGCGCTCGAGTGTCTCAAGGCGCTC 586  
QY CACTTCGCGCGCATGTTGTCAAGATTCGATCTTCAAGATCTCGCTGCTGACCGCGCT 712  
DB CACTTCGCGCGCATGTTGTCAAGATTCGATCTTCAAGATCTCGCTGCTGACCGCGCT 646  
QY CACGTGCTTCACGCTCTCAAGCTTCACAAAGTCTCCGCTTTCCACCGAGTCTTCTTTTC 772  
DB CACATGCTGCGAGAGCTCAAGCTTCGAAAGTCTCCGCTTTCTCCACCGAGTCTTCTTCG 706  
QY CATATCGGTGCGCTTTTGAAGAGTTTAAAGTCTTTTGGAGGAAAGCTCAATTCCT 832  
DB CTCGTGCGCGCTCTCTGCGAGTCACTGAGAACATTTATTTCTGGAGGAATGCTCAATTCCT 766  
QY GAGAGGACGAGAGATGCTACAGAGTCTTGAATTAATACAGTCTTCTGAGACTCTC 892  
DB GATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 826  
QY AATTTTCTTCTGACAGACATTCGCTGTTGGAAGATTTGAGGACCTTGAACCTTTTAGCTAA 952  
DB AAGTCAAGTCACTAATTTCAATGAAGATCACTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 943  
QY TTTCTTAAAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1072  
DB TTTCTTAAAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1003  
QY GA-----AAGATPACTCTGCTATATCATTTACCAGCAAGTTATGTCGATTTGGTGTAAACA 1126  
DB GAGCTCAGTAAAGTATGGAATTTAAATCCCTTCAAGCTGCTGCTTCTTAGGACTTACG 1063  
QY TATATTTGGAAGAAATGAGTGGCCATGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1186  
DB TACATGGGACAAACGAGATGCCATTAATCTTCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1123  
QY GATCTCTCTATGCAATGCTAGACGAGGATCATTTGATGTTAATCAAGAGTGTCTCA 1246  
DB GACTTGGAGTACACTTTCTTCCACCTGGAAGATCACTGCCAATCATTCGAAATGTTCC 1183  
QY AATCTGGAAGTCTCTGAGCAAGAAATGTAATTTGGAGATAGAGGTTAGAGTTTCTTGGT 1306  
DB AACTTACTAGTTCTTTCGGGTGAGGAATGTGATTGGAGATAGAGGATTTAGGGGTTTGTGCA 1243  
QY CGTTGTTGTAAGAGGCTTAAAGGCTTAGGATTTGAAAGGCGCATGATGATCAAGGAATG 1366  
DB GAGCTTGAATACATGCGCTGTTTATGTTGTTGATTAATAAATGCAATCTCTGGAACATATT 1486  
QY GAACTGGAATACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1423  
DB GGAATCTCTTGAAGAACCTCTGATTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCG 1546  
QY ATAACTGATTTGCAAAATCTTTTGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1483  
DB ATAACTGATTTGCAAAATCTTTTGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1606  
QY ATAACTGATTTGCAAAATCTTTTGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1543  
DB ATAACTGATTTGCAAAATCTTTTGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1666  
QY ATAACTGATTTGCAAAATCTTTTGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG

TITLE	Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
JOURNAL	Science 301 (5631), 376-379 (2003)
MEDLINE	22752273
PUBMED	12869764
REFERENCE	2 (bases 1 to 2300)
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujiyama, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hirooka, T., Horii, F., Hotte, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imorani, K., Ishibiki, J., Ishikawa, M., Itch, M., Kagawa, I., Morita, K., Nabiki, J., Shibui, Y.

38 TCGCTCTGCTACACACACACCCCGGCTGCCTCCGCCCGCTTCCCGCACCTCGAGTCG 532

**COMMENT**

## ORIGIN

298	DB	GCATGGCGTACTTCACACAGCCGCGCTGTTTCAGGCGGTTCCCTGCGCTCGAGTCG	357
533	QY	CTCAAGCTCAAGGGCAAGCCCGAGCGCAATGTTCAACTTGATACCGAGAGATTGGGCG	592
358	DB	CTCAAGCTCAAGGCCAAGCCCGGCGGCATGTTTCAACCTCATCCCGAGACTTGGGCG	417
593	QY	GGACAGTCACATCCCTGGGCTCAAGAGATTTCTCAGTACTCTCGATTGGCTCAAGAGCTC	652
418	DB	GGTCGCGCTCGCGGTGGATTCGCGAGCTCTCGGCCCTGTTTCACTTCTCAAGGGCGTCT	477
653	QY	CACCTCGCGCGCATGATTGTCAAGGATTCGATCTTTCAGAACTCTCGTCTGACCGCGGT	712
478	DB	CACCTCGCGGAGTATCTGTCCGACGACGACCTCGACGTCTCTGTCCGCGCCAAAGCG	537
713	QY	CAGTGTCTTCA CGCTCTCAAGCTTGA CAAGTGTCCGGTTTCA CCACCGATGCTCTTTTC	772
538	DB	CACATGCTCTCCCTCGTTCAGGCTTGACCGCTGCTCTGGCTTCTTCAACATCTCCCTCGCC	597
773	QY	CATATCGGTCGCTTTTGCAGAGTTTTAAGAGTCTTGTCTTTTGGAGAAAGCTCAATCTT	832
598	DB	CTCGTCCCGCACCTGCAAGAACTTTGAACGTTATTCCTGGAGGATAGCATAAATTGCT	657
833	QY	GAGAAGS--ACGGAGAATGGCTTACACGAGCTTCCTTGAATTAATACAGTTCTTGAGACT	889
658	DB	GAGAAAGAAATGACCAATGGAATTCGGAGCTTGTACCAACAATATGTCTCTTGAGACA	717
890	QY	CTCAATTTTTACTTTGACAGACATTTGCTGTGTGAAGATTGAGGACCTTGAATCTTTAGCT	949
718	DB	CTGAATTTCTTTTGTGACAGATCTCAGGCGATCCCCA---GCATATCTTACCTCTTGTG	774
950	QY	AAAAATGGCCCCAACTTAGTGTCTGTAACCTTACTGACTGTGGAATACTGGATCTTGTG	1009
775	DB	CGCAATGTGCGAAGGCTGAAAGTTTTGAAGATTAGTGAAGTCTTCACTGTGACCTGGTT	834
1010	QY	AACTTTTAAAGCATGCCTCGCGCTGGAGAGTTTTGTGGAGGCACTTACACGAGGAA	1069
835	DB	GATTTGTTCCGTTACAGCAGAAATATTACAGACTTTTGGCGTGTTCCCTTGTATGATCAG	894
1070	QY	CCGAAAGATACTCTGTATA-----TCATTACCGACAAAGTTATGTGCA	1114
895	DB	GGTCNAGTTGAAGAGTAGAATATGAAAACTACTATTTTCCCTCTTATTGCTCCGC	954
1115	QY	TTGGTTTACATATATTGGAAGATGAGTTGCGCCATGTGTTCATGTTTCAGCGCTA	1174
955	DB	TTGAGTTTGTCTACATATGGGAAACAAAGAGATGCAAGTGTATTTCCATATGTGTGCA	1014
1175	QY	CTAAAAAATTTGATCTCCTTATGCAATGCTAGACACGAGGATCATTTATATGTTAAATC	1234
1015	DB	CTCAGAAGTTGATCTTCAATTCACATTCCTTTCCACAGAGATCACTGTCAGCTGGTC	1074
1235	QY	CAAAAGTGTCCAAATCTGGAAGTCTTTGACACAAGGAATGTAATTGAGATAGAGGGTTA	1294
1075	DB	CAGCGCTGCCAAATCTGGAATCTTAGAGGTGAGGGATGTCATAGGGGATCTGTGGTTA	1134
1295	QY	GAGTTCCTCGTCTGTGTGAGAGGCTAAAAAGGCTTAGGATTGAAAGGGCGCATGAT	1354
1135	DB	GAGTTGTGTGACAAACCTGCAAGAAATTCAGAGACTTCGAGTAGAGAGAGAGATGAT	1194
1355	QY	GATCAAGGA---ATGGAGGATGAAGAAGTACTGTGTCCCATAGAGGCTAATAGCGTTG	1411
1195	DB	GATCAAGGAGTCTTGAGGATGAACATGTTATGTTTACACAGGTGGGTTTGTATGGCTGA	1254
1412	QY	TCACAGGCTGTTTCAGAGCTTGAATACATGGCTGTTTATGTGCTGATATTCAGAAATGCA	1471
1255	DB	GCTCAAGGCTGCCCTCAATTTAGAAATATCGGGCAGTGCAATGTCACATGACATACAAATGCT	1314
1472	QY	TCTCTGGAACATATTGGAACCTCACTTGAAGAACCTCTGTGATTTTCGCTTGTCTGCTT	1531
1315	DB	GCTCTAGAAGCCATCGGTACATACAGCAGCAGTCTTAAAGATTTCCGGTCTGCTCTGCT	1374
1532	QY	GACCATGAAGAGAGATAACTGATTTTGCCACTTGACAAATGGGGTGAGGCTCTACTGAGG	1591
1375	DB	GATAGAGAAGCAAAATATTAACAGAAATCACTCCCTTGCAATGGGGTGGCTTTACTGAGA	1434

RESULT 11  
AK100694

LOCUS  
DEFINITION

ACCESSION

VERSION  
KEYWORDS

**SOURCE**  
**ORGANISATION**

511

## REFERENCE

## AUTHORS

12869764  
2 (bases 1 to 2924)  
Adachi J., Aizawa K., Akimura T., Arakawa T., Carninci P., Doi K.,  
Fujimura T., Fukuda S., Hanagaki T., Hara A., Hashizume W.,  
Hayashida K., Hayashizaki Y., Hayatsu N., Hiramoto K., Hiraoka T.,  
Hori F., Hotta I., Iida J., Iida Y., Ikeda R., Imamura K.,  
Imotani K., Ishibiki J., Ishii Y., Ishikawa M., Itoh M., Kagawa I.,  
Kanagawa S., Katoh H., Kawagashira N., Kawai J., Kawamata M.,  
Kikuchi S., Kishikawa-Hirozane T., Kishimoto N., Kobayashi M.,  
Kodama T., Kojima K., Kojima Y., Kondo S., Konno H., Kouda M.,  
Koya S., Kurihara C., Kurosaki T., Kusumegi T., Li C., Lu M.,  
Maeda H., Matsubara C., Matsuyama T., Miura J., Miyazaki A.,  
Mizuno K., Murakami K., Murata M., Nagata T., Nakamura M.,  
Namiki T., Narikawa R., Niikura J., Nishi K., Nomura K.,  
Numasaki R., Ohneda E., Ohno M., Ohtsuki K., Oka M., Ooka H.,  
Ootani N., Ota Y., Otomo Y., Ryu R., Saitoh H., Sakai C., Sakai K.,  
Sakazume N., Sano H., Sasaki D., Sato K., Satoh K., Shibata K.,  
Shinagawa A., Shiraki T., Shishiki T., Sogabe Y., Sugano S.,  
Sugiyama A., Suzuki K., Suzuki Y., Tagami M., Tagami-Takeda Y.,  
Tagawa A., Takahashi F., Takaku-Akahira S., Tanaka T., Tomaru A.,  
Toya T., Tsunoda Y., Ueda M., Waki K., Xie Q., Yahagi W.,  
Yamada H., Yamamoto M., Yasunishi A., Yazaki J., Yokomizo S. and  
Yoshimura A.

**TITLE**  
Direct Submission

**JOURNAL**  
Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of  
Agrobiological Sciences, Department of Molecular Genetics, Head of  
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki  
305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp,  
Tel:81-29-338-7007, Fax:81-29-838-7007)

**COMMENT**  
This clone is one of the 28K full-length cDNA clones from japonica  
rice.  
URL : <http://cdna01.dna.affrc.go.jp/cdna/>  
NIAS Rice Full-Length cDNA Project Team: Kikuchi S., Satoh K.,  
Yamagata T., Kawagashira N., Doi K., Kishimoto N., Yazaki J.,  
Ishikawa M., Yamada H., Ooka H., Hotta I., Kojima K., Mamiki T.,  
Ohneda E., Yahagi W., Suzuki K., Li C., Ohtsuki K., Shishiki T. and  
Yamamoto M.

**FEATURES**  
FAIS Genome Sequencing & Analysis Group: Otomo Y., Iida Y.,  
Fujimura T., Ikeda R., Ishibiki J., Kawamata M., Kobayashi M.,  
Kodama T., Kurosaki T., Kusumegi T., Lu M., Masuda H., Miura J.,  
Mizuno K., Narikawa R., Niikura J., Oka M., Ryu R., Sugano S.,  
Sugiyama A., Suzuki Y., Tsunoda Y., Ueda M., Xie Q., Yokomizo S.,  
Yoshimura A., Matsubara C. and Murakami K.

**source**  
Genome Exploration Research Group in Riken Genomic Sciences Center  
and Genome Science Laboratory in Riken: Adachi J., Aizawa K.,  
Akimura T., Arakawa T., Carninci P., Fukuda S., Hanagaki T.,  
Hara A., Hashizume W., Hayashida K., Hayatsu N., Hiramoto K.,  
Hiraoka T., Hori F., Iida J., Imamura K., Imotani K., Ishii Y.,  
Itoh M., Kagawa I., Kanagawa S., Katoh H., Kawai J.,  
Kishikawa-Hirozane T., Kojima Y., Kondo S., Konno H., Kouda M.,  
Koya S., Kurihara C., Matsuyama T., Miyazaki A., Murata M.,  
Nakamura M., Nishi K., Nomura K., Numasaki R., Ohno M., Otsato N.,  
Ota Y., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
Sasaki D., Sato K., Shibata K., Shinagawa A., Shiraki T.,  
Sogabe Y., Tagami M., Tagami-Takeda Y., Tagawa A., Takahashi F.,  
Takaku-Akahira S., Tanaka T., Tomaru A., Toya T., Waki K.,  
Yasunishi A. and Hayashizaki Y.

**ORIGIN**  
Location/Qualifiers  
1. 2924  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="J023114118"

**Query Match**  
Query Match 25.4%; Score 581.6; DB 8; Length 2924;  
Best Local Similarity 61.1%; Pred. No. 1.7e-126;  
Matches 1070; Conservative 0; Mismatches 629; Indels 51; Gaps 6

**Qy** 353 GACGTGGTCTCGACTCGGTCATCCCTTACATCGACGACCCGACGACCGCCTT 412  
Db 761 GACCTTGGCGTGGGCTGCTGATGGGTTCTGTGGAGGACCCCTGGGACCGGACGACCATC 820

1472 TCCTGGAACATATTGGAACTCACTTGAAGACCTCTGTGATTTTCGCTTGGCTTGTGCTT 1531  
 1898 GCTCTAGAACCCATCGGTATACATACAGACAGCTTTAAGATTTCCGGTTTCTCTGCTT 1957  
 1532 GACCATGAGAGAGAACTACTGATTTGACATTTGACATTTGAGGCTGAGGCTCTACTGAGG 1591  
 1958 GATAGAGAGAGAAATATACAGATACACCCCTTGACATTTGAGGCTGAGGCTCTACTGAGG 2017  
 1592 GCTGTGACAACTGAGGAGATTGCTCTATATCTCAGCGGTGCGGGTGTGATGATGA 1651  
 2018 GCTGTGACCAAGCTCCGAGGTGCGATTTATGTAAGACCTGCGGCTCTATCAGATGT 2077  
 1652 GCTGTGATCATTTGAGGAGCAATACAGTCCAAATGTGAGATGATGCTGCTGTTATG 1711  
 2078 GCTGTGAGGATGATGCTGAGGATTTGATGATGATGATGATGATGATGATGATGATG 2137  
 1712 GGGGAGTCTGATGACGAGGCTTTTGGAGTTGCTTAAGGCTGCTTCTAGTCTTCAAGAACTT 1771  
 2138 GGGGAGTCTGACCAAGGACTGCTTCAACTCTCAACAGGATGCCAAGCTTGACAAATG 2197  
 1772 GAATGAGAGGCTGTTATTTTCACTGACGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTG 1831  
 2198 GAGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2254  
 1832 ACTTCTCTTAGGTACTTGTGGGTGCAAGGTTATGTTATCTCTCACTCTGACGCTGATCTT 1891  
 2255 AATCACTGATACCTATGAGGCTGAGGATGATGATGATGATGATGATGATGATGATGAT 2314  
 1892 TTGTTAATGCTGCTGACCTTTTGGAACTGATGATGATGATGATGATGATGATGATGATG 1951  
 2315 ATGCAATGCTGCGACCTTCTGGAACATGATGATGATGATGATGATGATGATGATGATG 2360  
 1952 AATCAATCCAGATGAGACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2011  
 2361 -----ATCAAGATGAAGTTTGCCTGAGGCTGAGGATGATGATGATGATGATGATGAT 2410  
 2012 TCCTTTCAGGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2071  
 2411 TCATTCAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2470  
 2072 TGGCTTGTATA 2081  
 2471 TGAGTGTATA 2480

RESULT 12  
 AF002109/c  
 LOCUS  
 DEFINITION Arabidopsis thaliana chromosome 2 clone T28M21 map C1C10A06, complete sequence.  
 ACCESSION AF002109  
 VERSION AF002109.3 GI:20198304  
 KEYWORDS HTG.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurooids II; Brassicales; Brassicaceae; Arabidopsis.  
 REFERENCE 1 (bases 1 to 108847)  
 AUTHORS Rounsley, S.D., Ketchum, K.A., Lin, X., Phillips, C.A., Brandon, R.C., Fuhrmann, J.L., White, O., Kerlavage, A.R., Adams, M.D., Somerville, C.R. and Venter, J.C.  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 108847)  
 AUTHORS Lin, X.  
 TITLE Direct Submission  
 JOURNAL Submitted (09-MAR-2000) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA  
 REFERENCE 3 (bases 1 to 108847)  
 AUTHORS Town, C.D. and Kaul, S.  
 TITLE Direct Submission  
 JOURNAL Submitted (27-FEB-2002) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA, cdtown@tigr.org

COMMENT On Apr 18, 2002 this sequence version replaced gi:6598791.

## FEATURES

## source

Location/Qualifiers

1..108847

/organism="Arabidopsis thaliana"

/mol\_type="genomic DNA"

/cultivar="Columbia"

/db\_xref="taxon:3702"

/chromosome="2"

/map="C1C10A06"

/clone="T28M21"

1..18752

/note="overlap with BAC clone T517

(AC003000:73873..92624)."

complement(2875..2905)

/rpt\_family="AT\_rich"

complement(5687..5709)

/rpt\_family="AT\_rich"

complement(8441..8512)

/gene="At2g39860"

complement(8441..8512)

/gene="At2g39860"

/product="tRNA-Asp"

12204..14036

/gene="At2g39870"

/note="synonym: T28M21.3; supported by cDNA:

gi:13877742 gb:AF370134.1 AF370134"

join(12204..12630,12715..13240,13330..13431,13520..14036)

/gene="At2g39870"

join(12358..12630,12715..13240,13330..13431,13520..13611)

/gene="At2g39870"

/codon\_start=1

/product="expressed protein"

/protein\_id="AA095272.1"

/db\_xref="GI:2088640"

/translation="MSRLPLPHLSVNLPTMKTSTSELTPRLGFENPEFPYDFDPSF

SGFTSPGDSSTEDDEDEFLAGTERLAPSTORLPSLKEEKKROVATSPQ

STLGSFGSNGSRSPILPFPATSSFRDNANDVISAAGEVARLKLGSVEPHL

PIQTESLLRQNAIAHAEHQRLIEQMLCSAQSDVSKLSENIRPRVNBEGLEN

PYRVNPNFTLPPQQAAPLKPAGTGVFLPRYPSPADSDSLKTPVTNPAILOPK

VFPNLNFDFTNIVGPRRSQDFECMLARSTVLARQGNFRAVSGGGLPQDMWY"

complement(13790..13823)

/rpt\_family="(TA)n"

14264..14329

/rpt\_family="(GAAA)n"

complement(16606..18140)

/gene="At2g39880"

/note="synonym: T28M21.4; supported by cDNA:

gi:5823308 gb:AF175988.1 AF175988"

complement(join(16606..17390,17822..18140))

/gene="At2g39880"

complement(join(16606..17390,17822..18140))

/gene="At2g39880"

/codon\_start=1

/product="putative MYB family transcription factor"

/protein\_id="AA095273.1"

/db\_xref="GI:2088641"

/translation="MNGEISRPPELISSNPKSPENAIHKAELAEIAKSDANGG

GSKVGPWLQDEALTLVQCGPNWNLISRGTPGRSGSKSLRWNCQDPLIKR

KPSDEEHMIMSAQVLGNKWSVIAKLPGRTDNAIKHWNLSNLRKPAEQWKIPLL

MNTETVYOLYPSMVTIRISNAPKEHLPOBEETGVLSDDKMDDEAPEQNSKTGV

YRVAMGAFSVCKPGYMAPCEGPLYQASRPDSLAGKFLQSLCVDPIIPSKGCGCN

HQDSTLSSSVIGSFVDEEHSALDKELLISINDLNTAWIRSGKEAEQSLKAD

DQFREYAHSKFGMVNNGVSGMVQDILRAUS"

complement(18576..18596)

/rpt\_family="AT\_rich"

complement(19036..19079)

/rpt\_family="(TA)n"

20589..20618

/rpt\_family="AT\_rich"

complement(21270..21305)

/rpt\_family="(TA)n"

21840..21868

/rpt\_family="AT\_rich"



```

QY 1797 GTGAAGCTGCACTGTGTGGCTGCAACACAAATGACTTCTCTAGGTACTTGTGGTGC 1856
Db 41786 GTGAGGGAGCAATCGTGCAGCGGTATCAAAATTCCTTCACTGAGATCTTGTGGGTAC 41727
QY 1857 AAGGTTATGGTATCTCCATCGACGCTGATCTTTTGGTAATGCTCGACCCCTTTTGA 1916
Db 41726 AAGGTTACAGAGCATCGATGACGGGACAAGATCTAATGACAGTGGCTAGACCGTACTGGA 41667
QY 1917 ACATTGAGTTGATCTTCTAGAAAGTGGCTACGATACCAATCCAGATGAGACTGTAG 1976
Db 41666 ACATCGAGCTGATTCATCAAGAAGAGTCCCGGAAGTGAATCAACAAGGAGAGATAAGAG 41607
QY 1977 TTGTTGAGCATCTCTGCTCATATCTTGTGCATATATTTCTCTGAGGGCAGAGATCAGATT 2036
Db 41606 AGATGAGGAGCATCGGCTCATATATTTGGCTTACTACTCTCTGCTGGCCAGAGACAGATT 41547
QY 2037 TTCCAGATCTGTTGTCCTTTGGACATGCGCACATGCGTTGATCCTAGAGCGGACAGAGC 2096
Db 41546 GTCCAAACACTGTTAGAGTCTGAGGAGGCCAATATGATG--ACCCAAAACAGAGTT 41489
QY 2097 TGTGTATATATACCAGTTTCTTTTGTGTTT 2126
Db 41488 TGTATATAAGATTTTGTAGTCTCGAGTTT 41459

RESULT 13
AX046779
LOCUS AX046779
DEFINITION Sequence 5 from Patent WO0068406.
ACCESSION AX046779
VERSION AX046779.1 GI:11876298
KEYWORDS Glycine max (soybean)
SOURCE Glycine max
ORGANISM Glycine max
REFERENCE 1
AUTHORS Caimi, P.G., Fanodu, O.O., Lee, J.M., Miao, G.H. and Maxwell, C.A.
TITLE Disease resistance factors
JOURNAL Patent: WO 0068406-A 5 16-NOV-2000;
E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES
Source 1. 482
/organism="Glycine max"
/mol_type="unassigned DNA"
/db_xref="taxon:3847"
ORIGIN
Query Match 18.0%; Score 412; DB 6; Length 482;
Best Local Similarity 97.7%; Pred. No. 1.8e-86;
Matches 471; Conservative 0; Mismatches 5; Indels 6; Gaps 5;

QY 334 GCGGAAGACACAGTGTGTGTCGACGTGGTCTCGACTGCGTATCCCTTACATCGACGACC 393
Db 1 GCGGAAGACACAGTGTGTGTCGACGTGGTCTCGACTGCGTATCCCTTACATCGACGACC 60
QY 394 CAAGGACCGGACGCGCGTTTCCAGGTGTGTGTCGACGTGTGTCGACGTGTGTCGACTCGCTCAC 453
Db 61 CAAGGACCGGACGCGCGTTTCCAGGTGTGTGTCGACGTGTGTCGACGTGTGTCGACTCGCTCAC 120
QY 454 CCGCAAGCAGTCAACATCGGCTCTGCTACACACACACACCGCGCTCGCTCGCGCGCG 513
Db 121 CCGCAAGCAGTCAACATCGGCTCTGCTACACACACACACCGCGCTCGCTCGCGCGCG 180
QY 514 CTTCCCGCAGCTCGAGTGTGCTCAAGCTCAAGGGGCAAGCCCGGAGCCGCAATGTTCAACTT 573
Db 181 CTTCCCGCAGCTCGAGTGTGCTCAAGCTCAAGGGGCAAGCCCGGAGCCGCAATGTTCAACTT 240
QY 574 GATACCCGAGGTTGGGGCGGACAGCTCACTCCCTGGGTCAAGAGATTTCTC-AGTACT 632
Db 241 GATACCCGAGGTTGGGGCGGACAGCTCACTCCCTGGGTCAAGAGATTTCTCAAGTACT 300

```

```

QY 633 TCGATTGCTCAAGAGCCTCCACTTCCGCCGATGATTGTCAA--GGATTCCGATCTTCAG 691
Db 301 TCGATTGCTCAAGAGCCTCCACTTCCGCCGATGATTGTCAAAGGATTTCGATCTTCAG 360
QY 692 AATCTCGTGTGACCGCGGTCACTGCTTCAAGTCTC--AAGCTTGAACAAGTGTCCGG 750
Db 361 AATCTCGTGTGACCGCGGTCACTGCTTCAAGTCTC--AAGCTTGAACAAGTGTCCGG 420
QY 751 TTTCCACACCGATGCTCTTTTCCCATATC--GTCGCTTTTGCAA--GAGTTTAAAGAGTCTT 807
Db 421 TTTCAACAACGATGCTCTTTTCCCATATCGGTCGCTTTTGCAAAGATTTTAAAGTCTT 480
QY 808 GT 809
Db 481 GT 482

```

```

RESULT 14
AP003279
LOCUS AP003279
DEFINITION Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,
ACCESSION AP003279 BA000010
VERSION AP003279.2 GI:15718426
KEYWORDS Oryza sativa (japonica cultivar-group)
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
REFERENCE 1
AUTHORS Sasaki, T., Matsumoto, T., Yamamoto, K., Sakata, K., Baba, T.,
Kataose, Y., Wu, J., Nilmura, Y., Cheng, Z., Nagamura, Y.,
Antonio, B.A., Kanamori, H., Hosokawa, S., Masukawa, M., Arikawa, K.,
Chiden, I., Hayashi, M., Okamoto, M., Ando, T., Aoki, H., Arita, K.,
Hamada, M., Harada, C., Hijishita, S., Honda, M., Ichikawa, Y.,
Idonuma, A., Iijima, M., Ikeda, M., Ikono, M., Itoh, S., Itoh, F.,
Itoh, Y., Itoh, Y., Iwabuchi, A., Kamiya, K., Karasawa, W., Katagiri, S.,
Kikuta, A., Kobayashi, N., Kono, I., Machita, K., Maehara, T.,
Mizuno, H., Mizubayashi, T., Mukai, Y., Nagasaki, H., Nakashima, M.,
Nakama, Y., Nakamichi, Y., Nakamura, M., Namiki, N., Negishi, M.,
Ohta, I., Ono, N., Saji, S., Sakai, K., Shibata, M., Shimokawa, T.,
Shomura, A., Song, J., Takazaki, Y., Terasawa, K., Tsuji, K., Waki, K.,
Yamagata, H., Yamane, H., Yoshiki, S., Yoshihara, R., Yukawa, K.,
Zhong, H., Iwama, H., Endo, T., Ito, H., Hahn, J.H., Kim, H.I., Eun, M.Y.,
Yano, M., Jiang, J. and Gojobori, T.
TITLE The genome sequence and structure of rice chromosome 1
JOURNAL Nature 420 (6913), 312-316 (2002)
MEDLINE 22337376
PUBMED 12447438
REFERENCE 2 (bases 1 to 134982)
AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE Direct Submission
JOURNAL Submitted (19-FEB-2001) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program; Kamondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)

```

```

COMMENT
On Sep 20, 2001 this sequence version replaced gi:13027309.
Genes were predicted from the integrated results of the following:
GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor
(October 1998 version). The genomic sequence was searched against
NCBI NonRedundant Protein database, nr
(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at
RGP. Protein homologies of the coding regions were searched against
NCBI NonRedundant Protein database with BLASTP2.0. ESTs represent
the identified cDNA sequences using BLASTN 2.0 with the
corresponding DDBJ accession no. and RGP clone ID.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative', and '-like protein'. A gene without
significant homology to any protein but with EST homology (covering

```



almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted with a gene prediction program is classified as a 'hypothetical' protein.  
 The orientation of the sequence is from T7 to SP6 of the PAC clone. This sequence of P0529E05 clone has an overlap with P0529H11 (DDBJ: AP004072) clone at the position 1 to 19,018 of 5' end. The sequence of this clone starts at the position 146,952 of P0529H11. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://rgp.dna.affrc.go.jp/GenomesSeq.html>.

## FEATURES

source	Location/Qualifiers
gene	<p>1.134982            /organism="Oryza sativa (japonica cultivar-group)"            /mol_type="genomic DNA"            /cultivar="Nipponbare"            /db_xref="taxon:39947"            /chromosome="1"            /clone="P0529E05"            complement(788..1069)            /gene="P0529E05.1"            complement(788..1069)            /gene="P0529E05.1"            /note="hypothetical protein"            /codon_start=1            /protein_id="BAB84386.1"            /db_xref="GI:18461193"            /translation="MVASVGGGGRVAAQPRRAGAVSATVAGAGGCGAWAERQWR QQDTNQVDEMTRDVRVELVTEKPKFLPVGNHPLVPGFPVVAIRD"</p>
CDS	<p>5288..5456,5562..5684,6335..6361))            /gene="P0529E05.2"            complement(join(3596..3889,3976..4274,4750..4839,5288..5456,5562..5684,6335..6361))            /gene="P0529E05.2"            complement(join(3596..3889,3976..4274,4750..4839,5288..5456,5562..5684,6335..6361))            /gene="P0529E05.2"            complement(join(3596..3889,3976..4274,4750..4839,5288..5456,5562..5684,6335..6361))            /note="contains EST R03279 (R3279)            hypothetical protein"            /codon_start=1            /protein_id="BAB84387.1"            /db_xref="GI:18461190"            /translation="MAMIFLLVQEI CMATSNFSEQRNIGLNGFTVYKLRDGSIIA VKRATNMVDRLSEFRSEIQTLSKEHLNLVKFLGLEHEDERLILVEYVNGSLR EHLDRGLRLEPESRLNTAIDIVHAVSYLHGVTDPHPI IHRDITKSNILLTDQLRAKY ADGFARLAPDNTAEATHTVMVKTAGYVDPYMRNTQLTDRSDVTSFGVLLVELLTG RPIERGRGRHLLTQWALRXCEDGDAVAMDARWRTSAVVAAMEKVMALAECTA PDRAAPAMRCAEVLWSIRDFPHQEQQRATAAAGARGKHDGTYGSPISLKE" complement(join(7920..8104,8190..8223,8332..8469,8893..9039,9136..9180,9283..9646,9758..9984,10424..10843)) /gene="P0529E05.3"            complement(join(7920..8104,8190..8223,8332..8469,8893..9039,9136..9180,9283..9646,9758..9984,10424..10843))            /gene="P0529E05.3"            complement(join(7920..8104,8190..8223,8332..8469,8893..9039,9136..9180,9283..9646,9758..9984,10424..10843))            /note="hypothetical protein"            /codon_start=1            /protein_id="BAB84388.1"            /db_xref="GI:18461191"            /translation="MAPPGQLPLASLLPLSAPPFVSGRRRLRLVLGRALPPPTW LPHGLSPALPPAPSRPPPPATSLPGAPGGGAQAQVLAFTVSPKVAW MLGLALCNADRVPMVAIVLSQAYGWTSPFAGVVSQSFELGVLSPILGALVDY YGGRVMAYGVALMSLTFSPAAARSLMLFLSTRVLLGAEVGLVPMNNVLRWF PRTSSAVGIMAGFOLGNTIGLLSPIIMSRAGIFGPFVIFGLFGLVNLWISAI SGTGEGNAQISAEHLDTITGOKLVKTQSGGERLRKVPFSLKSWPTWALISANAM HSWGYFVLSMWPYFKTIVHVNIREAAWFSALPVMVAVLYGAVGVSDRLIQNGTS ILLTRKIMQITGFGVPGVALLGLNAKSPVITASAWLITAVGLKSFHSGFLVNLQETA PQYAGVLHMGNTAGTFAAILGTGAGFFVDRMGSRFGFLITLSLIFSSTLFLWDIFA TGERVDFDGTG"</p>
gene	<p>11252..11596            /gene="P0529E05.4"            11252..11596            /gene="P0529E05.4"            /note="hypothetical protein"            /codon_start=1            /protein_id="BAB84389.1"            /db_xref="GI:18461192"</p>
CDS	<p>complement(join(35905..36225,36435..36510,38277..39001,42150..42599))            /gene="P0529E05.8"            complement(join(35905..36225,36435..36510,38277..39001,42150..42599))            /gene="P0529E05.8"</p>
gene	<p>complement(30042..34424)            /gene="P0529E05.7"            complement(30042..34424)            /gene="P0529E05.7"            /note="pseudogene, similar to TNP-like transposase element"</p>
misc_feature	<p>complement(30042..34424)            /gene="P0529E05.7"            complement(30042..34424)            /gene="P0529E05.7"            /note="pseudogene, similar to TNP-like transposase element"</p>
gene	<p>complement(join(35905..36225,36435..36510,38277..39001,42150..42599))            /gene="P0529E05.8"            complement(join(35905..36225,36435..36510,38277..39001,42150..42599))            /gene="P0529E05.8"</p>
CDS	<p>complement(join(35905..36225,36435..36510,38277..39001,42150..42599))            /gene="P0529E05.8"</p>
gene	<p>complement(join(14141..14355,14614..14942,15117..15432,15672..15826,15966..16014,16692..18537))            /gene="P0529E05.5"            complement(join(14141..14355,14614..14942,15117..15432,15672..15826,15966..16014,16692..18537))            /gene="P0529E05.5"            /note="hypothetical protein"            /codon_start=1            /protein_id="BAB84390.1"            /db_xref="GI:18461193"            /translation="MTTLQLDPPVVPARRSAATSCDLHPDETFTGCTACLERLAGL EASAAAAAGPKSTSAIRSLFARFPFGAGGSSVAGAGSLDPLRRKSCSCGRRGHQ ALAAAAAAGATARADPQRSCDVRGHTLSLFLQDRGRVSSSAADAPPHQ OPPPPPPPTPDDFLDDEI PVMEHDEIMFVFEVIVVDTSGETEPNVVAREGA MKDHMPSSOPKPPPTKDLAESFWAASVSKWKWRKQKLKKEAASVAAEAM PPEKPSFELRRRLRGEAGSELAGRRSCDTPRSLDAGRMSVDDAGFSWDEPR ASWDGYLFGAGLGRAPPLSLRPPMWSVLETPATIVERSGOIPEVEDADIDP PGSLQTRDYLDSTRERSSVRSPFSEVTEPKPATITANGKESPLGSEF YHFHAEDLLDRGFSSNLIEDISASLEALSGPGSAKPRWRKAMSLWGFHRRAS GRSGSPSDIADRSFSEAWPDLRVYSPGRVDMGLRFLYTPMRSGGGGGRARVAAA VDVNGSSARRREHVLERNRSPGRVDMGLRFLYTPMRSGGGGGRARVAAA SETVKGTRRLALIVMRRLCLPLRLVDLWSSADADTASATWRPBRRCVLSDELL ASIAPVLAAGHGNRPTQERTDSGIDRNDGKDKHETAGOSSAPASAPAKNHDEP AVASPVGRQKRGHFRRAELGSTRVQVMPHARCPPEPGFNARLATGRNVHVFRR TSSRVPGLQHGMLRQSETGHIRRHLINACCFVSLPGVAKRFTCNQSDITKTK OTSRTNPFSSMRGALILDQMARANTHRTIPTCAEDMQMHVMSCHMCTFISHG STVRNADMAIYLPRAKVDVETEDTTSVRPCHFSVLTKLSIKHLICRSKW1" complement(join(23661..23843,23923..24021,24116..24188,24273..24383,24478..24651,24777..25855,25951..27372)) /gene="P0529E05.6"            complement(join(23661..23843,23923..24021,24116..24188,24273..24383,24478..24651,24777..25855,25951..27372))            /gene="P0529E05.6"            complement(join(23661..23843,23923..24021,24116..24188,24273..24383,24478..24651,24777..25855,25951..27372))            /note="contains EST AU090537 (C12241)            unknown protein"            /codon_start=1            /protein_id="BAB84391.1"            /db_xref="GI:18461194"            /translation="MADRDEOILYDTIAEGSSVYNEEGNKDPQDLNEEGNVERD AGNOGHMERDVEGQNEEASGSSQSVQKARGORGAKLEGHHITTEVEDRGP SAPAAEKYVRHSGVVRNDVNFVSWRRTRGDEHSFVPSDEKMLTMTLTFVAY TLPAGTEDKVRMTLKKMAEQFSKGLYKTLIKQTPNFTFFKLHWHQDFVAY KTGEOQQAAMERKENAKKYHHHLSGGSVSVMPKMEEMASLLERGLEPATANWE ERSKFYWAHGGTLNADGSLVFGDQIREAARLTNAVEASSQGTFRDRKDELSLA LQTPHPGTRGVIPWKIGFKEDITHVSQMRSKRDTAKIADLEFRSSYELSMQ EEVARKVDERMAHRSNNPQPTIPPAVSPSGNRSSCASTGVQSGSDMAQVQDEST CFVDDITQTTCELHILPKNLSIKVASGNAIPDPSGTVHCRPIPVGYSVVELVSG AYEDLDPYGGDGTDRDTSAILMHRRIYILPRQOASRAPSPAPSPQDPA PSHPHAPATPQAPATPPRSPPTPRSPPTSPQAPAFPSKRAPQAPPAHTR ATTKKAIDAAKNDPGVCTQBEDAYVASVEKQKPRSPKPIPIDPSVMAFPMGM SAPAKAIKLSYVERTLKASGKSPVLPQEQNPQIEPLVTGKEMMTKEFTIDTG LTDDQLGVTPIEAEKVYKELGKPLVKPELLOFLPTOMYKPHLYMENSATGREMI GARIDMDLQGGDDILWIPKFGIYQLDLDVIMEIORARRRGVDFDTGIDARKV NVAMLDQYQATDNLVHLLKAHYKTFILLPNTBFWHLLLDLSACTVNVYDMD KKESTDFKVFELDRWYFRHLVRGKWRRLRRKFPCKAKQKQLNDLGVYVCVC HCLADQIITRELDQIFRMRDLNLTTHKEFTAVQEQLMFNEQLDPRGEFYDGNII HRSLSAELATITTSKS"</p>
gene	<p>complement(30042..34424)            /gene="P0529E05.7"            complement(30042..34424)            /gene="P0529E05.7"            /note="pseudogene, similar to TNP-like transposase element"</p>
misc_feature	<p>complement(30042..34424)            /gene="P0529E05.7"            complement(30042..34424)            /gene="P0529E05.7"            /note="pseudogene, similar to TNP-like transposase element"</p>
gene	<p>complement(join(35905..36225,36435..36510,38277..39001,42150..42599))            /gene="P0529E05.8"            complement(join(35905..36225,36435..36510,38277..39001,42150..42599))            /gene="P0529E05.8"</p>
CDS	<p>complement(join(35905..36225,36435..36510,38277..39001,42150..42599))            /gene="P0529E05.8"</p>

/note="hypothetical protein"

Query Match 16.5%; Score 376.8; DB 8; Length 134982;  
Best Local Similarity 58.8%; Pred. No. 9,7e-78;  
Matches 802; Conservative 0; Mismatches 467; Indels 95; Gaps 5;

QY 798 TAAGAGCTCTGTTTGGAGGAAGCTCAATCTTGAGAGGACGGAGAAATGGCTACACG 857  
Db 68849 TGAGAACTTATTTCTGGAGAGTGCCATTAATCTGATAGGGTGGTGAATGCTTCATG 68908  
QY 858 AGCTTGCTTTGAATAATACAGTCTTTGAGACTCTCAATTTTACTTGACAGACATGCTG 917  
Db 68909 AACTTGCTGCAAAATCTGTTCTGGTGACACTGAACTTCTACATGACTGAACT---CA 68965  
QY 918 TTGTGAAGATTGAGGACCTTGAATTTAGCTAAATTTGCCCAACTAGTCTGTGA 977  
Db 68966 AAGTGGCGCAGCTGATCTAGAGCTTCTTCAAGAATTTGCAAGTCAATGATTTATTGA 69025  
QY 978 AACTTACTGACTGTGAATACTGGATCTTGTGAACCTTTTAAAGCATGCCCTCTGCGCTGG 1037  
Db 69026 AGATGAGTGAGTGATCTTTCAGATCTGATAGTTTCTTCAACAGCAATGGCTGC 69085  
QY 1038 AAGAGTTTGTGAGGACCACTAACAGGAACACGA-----AAGATCTCTGCTATAT 1091  
Db 69086 AAGACTTTGTGGAGGAGCATCTTACGAGGTAGGAGAGCTCACCAAGTATGAAAAGTTA 69145  
QY 1092 CATTAACGACAAAGTTATGCTGATTTGGTTTAAACATATATGGAAGAATGAGTTGCCA 1151  
Db 69146 AGTTCCACCCAGCAATATGCTTCTTGGGGCTTACTACATGGGAACAATGAGATGCTG 69205  
QY 1152 TTGTGTTTCATGTTTGCAGCCGCTACTAAATAATTTGGATCTCTCTATGCAATGCTAGACA 1211  
Db 69206 TTATCTTCCCTTTTCGATGAACTCAAGAACTGGACTTGCATACACITTTTCTCAA 69265  
QY 1212 CGGAGATCATATGATGATTAATCCAAAGGTGTCAAATCTGGAAGTCTTGAGACA---- 1267  
Db 69266 CAGAAGATCATTTGTCAGATTTATGCAAAATGTCACCAATCTACTAATTTCTTGAGGTAATAT 69325  
QY 1268 ----- 1267  
Db 69326 TTCTGTATACAGCATTTGATTAATGTTGTTGAATTAGTATATGCAATATTAACAAG 69385  
QY 1268 -----AGGAATGTAATTTGGAGATAGAGGGTTAGAGGTTCTTGCTGCTG 1311  
Db 69386 ATGTCTCTCAACAGGTGAGAACGTCATAGGAGATAGAGGGCTAGAAATTTGCTGTATAC 69445  
QY 1312 TTGTAAGAGGCTAAAGGCTTAGATTTGAAGGGGCGATGATCATCAAGAAATGGAGGA 1371  
Db 69446 ATGCAAGAAGCTACGAAGACTCCGAATTTGAGCGGGTGATGATGATCCAGGTCTGACGA 69505  
QY 1372 TGAAGAAGGTACTGTGCTCCCATAGAGGCTTAATAGCCTTTGTACAGGGCTCTTCAGAGCT 1431  
Db 69506 AGAGCAAGGAGAGTTTCTAGTAGGCTTGACAGCCGTTGCTGTTGTTCCGCTGAAT 69565  
QY 1432 TGAATACATGCTGTTATGCTGTGATATTAATAATGCATCTCTGGAACATATTTGAAC 1491  
Db 69566 GGAGTACATAGCTGCTTATGATCGATATCACCAATGGGCGCTGAGTCTATTTGGGAC 69625  
QY 1492 TCACCTTGAAGACCTCTGATTTTCGCTTGTGTTGCTTGCACCATGAAGAGAGATAAC 1551  
Db 69626 TTTCTGCAAAAATCTATACGACTTTTCGCTTGTGCTACTTGACAGAGAAAGACAGGTAAAC 69685  
QY 1552 TGATTTGCCACTTGCAATGGGCTGAGGGCTCTACTGAGGGGCTGTGCAAGCTGAGGAG 1611  
Db 69686 AGATCTGCCACTTGCAATGGTGTCTGCTCTGTATAGAAATTTGCAAAAGCTTCGGAG 69745  
QY 1612 ATTTGCTCTATATCTCAGCGTGGCGGTTGATCTGATGAGCCCTTGGTTACATTTGACA 1671  
Db 69746 GTTTGCTCTCTACCTTTAGACCAGGAGGCTTTTCAGATGATGCTTACTATCATCGACA 69805  
QY 1672 ATACAGTCCAAATGTGAGATGCTGCTGTTGGTTATGTTGGGGAGTCTGATGAGGGCT 1731  
Db 69806 GTACAGTGGAAATATCCATATATGCTACTGGGCAATGTTGGTGAATCTGACCATGGAT 69865

QY 1732 TTTGGAGTTCCGTAAGGGCTGTCTAGTCTTTCAGAACTTGAATGAGAGGGTGTATT 1791  
Db 69866 GATCCGTTTCGAGTGGGCTGCACCAACTTTCAGAGCTTGAATGAGAGCTG---CTG 69922  
QY 1792 TTTCAAGTGAACGTCGACTTTGCTGTGGCTGCAACAAATGCACTTCTTTAGGTACTTGTG 1851  
Db 69923 CTTTACGAGCGAGCTTTTGTCCCTGCTTACTGTCAGATGCCCTCCCTGAGATACATATG 69982  
QY 1852 GGTGCAAGGTTATGGTGATCTCCATCTGCACTGATCTTTTGGTAAATGGCTCGACCCCT 1911  
Db 69983 GGTGCAAGGATACAGAGCATCTCAACAGCCCTTGACCTTCTCATGGCCAGGCTTT 70042  
QY 1912 TTGGAACTATTGAGTTGATTTCTTTCTAGAAGGTGGCTACGAATACCAATCCAGATGAGAC 1971  
Db 70043 CTGGAACATCGAGTTTACACCTCCGAGCCCTGAGAGCTTTTAATCATATGACAGAAGATGG 70102  
QY 1972 TGTA-----GTTGTTGACATCTGCTCATATTTCTTGCATATTTCTTCTGACGGCAGAG 2028  
Db 70103 AGAACCTGTGTGATAGCATGCTCAGGTTCTTGTCTACTATTCCTTGTCTGGAAGGAG 70162  
QY 2029 ATCAGATTTTCCAGATACTGTTGTGCTTTTGGACACTGGCCACAT 2072  
Db 70163 GTCTGACTGCCCTCAGTGGGTGATCCCTTGATCTCTGCTGAT 70206

RESULT 15  
AX046791  
LOCUS 1728 bp DNA linear PAT 15-DEC-2000  
DEFINITION Sequence 17 from Patent WO068406.  
ACCESSION AX046791  
VERSION AX046791.1 GI:11876304  
KEYWORDS  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD  
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1  
Caimi, P.G., Famodu, O.O., Lee, J.M., Miao, G.H. and Maxwell, C.A.  
Disease resistance factors  
Patent: WO 0068406-A 17 NOV-2000;  
E.I. DU PONT DE NEMOURS AND COMPANY (US)  
Location/Qualifiers  
FEATURES  
source 1..1728  
/organism="Zea mays"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4577"

ORIGIN

Query Match 16.3%; Score 372.4; DB 6; Length 1728;  
Best Local Similarity 61.0%; Pred. No. 5.3e-77;  
Matches 702; Conservative 0; Mismatches 421; Indels 27; Gaps 5;

QY 807 TGTGTTTGGAGGAAGCTCAATTTCTTGAGAGG---ACGGAGATGCTACACGAGCTTG 863  
Db 34 TTTTCTCTGGAAGAAGCACAAATTTGATGAGAAGAAAATGATGATGATCGTGGCTTG 93  
QY 864 CTTTGAATAATACAGTTCTTGAGACTCTCAATTTTACTTTTACGACAGACATTTCTGTTGA 923  
Db 94 CTACAGCAATCTGTTCTTGAGACACTGAAATTTCTTAAACAGATCTCAGGCGATCCC 153  
QY 924 AGATTGAGGACCTTGAACTTTTGTAGCTAAATAATGCCCAACTTGTAGTCTGTGAACTTA 983  
Db 154 CA---GAGTATCTTACCTCTCTGTCGCAACTGTCAACGATTGAAACTCTGAAGATTA 210  
QY 984 CTGACTGGAATACTGATCTTGTGACTTCTTTAGCATGCTCTGCGCTGGAAGAGT 1043  
Db 211 GTGAATGTTTCATGCCCGATCTGCTCAGTTTGTTCGAACTGCAACAACTACAGAGT 270  
QY 1044 TTTCTGAGGAGCACCTCAACAGGAGAACAGAAAGATATCTGTCTATA----- 1090  
Db 271 TCGCTGTGGTTCCTTTTGAAGAGCAGGCTCAACCTGTGGCAAGTAGAAATATGAGAACT 330

Search completed: April 19, 2004, 22:56:44  
Job time : 8819 secs

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 18:00:14 ; Search time 875 Seconds  
(without alignments)  
11108.431 Million cell updates/sec

Title: US-10-009-791-21  
Perfect score: 2288  
Sequence: 1 gcacgagccacacgttaca.....acattctcaattttgtgaat 2288

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues  
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq 29Jan04.\*  
1: Geneseqn1980s.\*  
2: Geneseqn1990s.\*  
3: Geneseqn2000s.\*  
4: Geneseqn2001as.\*  
5: Geneseqn2001bs.\*  
6: Geneseqn2002as.\*  
7: Geneseqn2003as.\*  
8: Geneseqn2003bs.\*  
9: Geneseqn2003cs.\*  
10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2288	100.0	2288	3	AAA95063 cDNA from
2	668.4	29.2	2240	3	AAA95062 cDNA from
3	412	18.0	482	3	AAN02373 COI1 clon
4	412	18.0	482	3	AAA95055 cDNA from
5	372.4	16.3	1728	3	AAA95061 cDNA from
6	296.6	13.0	844	3	AAN02372 COI1 cont
7	296.6	13.0	844	3	AAA95054 cDNA from
8	233.2	10.2	1074	3	AAA95066 cDNA from
9	227.2	9.9	719	6	ABQ65584 Arabidops
10	223.4	9.8	794	3	AAA95056 cDNA from
11	223.4	9.8	794	3	AAN02374 COI1 clon
12	193.8	8.5	537	5	AAN02373 Peppermin
13	193.8	8.5	537	5	AAN02373 Peppermin
14	131.6	5.8	1392	8	ADA49036 Wheat gen
15	129	5.6	701	3	AAA95053 cDNA from
16	124.4	5.4	283	6	ABL72195 Corn tass
17	114.8	5.0	562	3	AAN02378 COI1 clon
18	114.8	5.0	562	3	AAA95060 cDNA from
19	107.8	4.7	486	3	AAA95065 cDNA from
20	105.2	4.6	266	7	ABX83406 Corn ear-
21	101.6	4.4	2361	3	AAC41105 Arabidops
22	96.2	4.2	1909	3	AAC42586 Arabidops
23	76.4	3.3	609	7	ABZ73247 Rice leaf

ALIGNMENTS

RESULT 1  
AAA95063  
ID AAA95063 standard; cDNA; 2288 BP.  
XX  
AC AAA95063;  
XX  
DT 12-FEB-2001 (first entry)  
DE  
DE cDNA from the soybean clone sgs4c.pk003.k23:fls homologous to COI1.  
XX  
KW Corn; rice; soybean; wheat; ATHCOR1; LLS1; disease resistance;  
KW A. thaliana coronatine-induced; COI1; herbicide; EST;  
KW expressed sequence tag; sgs4c.pk003.k23:fls; 88.  
XX  
OS Glycine max.  
XX  
FH Key Location/Qualifiers  
FT CDS 266..2086  
FT /\*tag= a  
FT /product= "COI1"  
FT /partial  
XX  
XX WO200068406-A2.  
XX  
XX 16-NOV-2000.  
XX  
XX 03-MAY-2000; 2000WO-US011955.  
XX  
XX 07-MAY-1999; 99US-0133041P.  
XX  
XX (DUPO ) DU PONT DE NEMOURS & CO E. I.  
XX  
XX Caimi PG, Fancdu OO, Lee J, Miao G, Maxwell CA;  
XX  
XX WPI: 2000-687649/57.  
XX  
XX P-PSDB; AAB23458.  
XX  
XX New nucleic acid sequences encoding new disease resistance factors, and to  
XX useful for producing plants with increased resistance to pathogens and to  
XX screen for herbicides.  
XX  
XX Claim 2; Page 59-60; 74pp; English.  
XX  
XX Expressed sequence tags (ESTs) from rice, wheat, soybean and corn cDNA  
XX libraries were found to be similar to cDNA encoding the Arabidopsis  
XX thaliana COI1 protein and the Zea mays LLS1 protein. The present sequence

24 67.6 3.0 577 3 AAA95064 cDNA from  
25 49.8 2.2 2000 7 ADA71938  
26 48 2.1 660 2 AAV50489  
27 48 2.1 7193 2 AAV50431  
28 48 2.1 7193 9 AAD61763  
29 46.8 2.0 2000 7 ADA71938  
30 46.8 2.0 3438 2 AAV39080  
31 46 2.0 6019 6 ABK31191  
32 46 2.0 6019 6 ABL70160  
33 46 2.0 6019 6 AAS61104  
34 45.4 2.0 17580 6 ABL54312  
35 44.8 2.0 17580 6 ABN80013  
36 44.8 2.0 1020 3 AAA30718  
37 44.8 2.0 1020 9 ADC22700  
38 44.8 2.0 7327 6 ABL33198  
39 44.4 1.9 113193 7 AAD54645  
40 44.2 1.9 6085 2 AAT70153  
41 44.2 1.9 34094 2 AAZ30163  
42 44.2 1.9 34094 8 ACC78130  
43 44 1.9 2000 7 ADA72854  
44 43.6 1.9 732 3 AAA39825  
45 43.4 1.9 1020 3 AAA30601

CC is cDNA from the soybean clone sgs4c.pk003.k23.fis which is homologous to  
 CC COI1. The COI1 and LSI1 proteins are involved in disease resistance. The  
 CC COI1 and LSI1 DNAs of the invention may be used to alter the expression  
 CC of COI1 and LSI1 protein in cells, particularly to produce transgenic  
 CC plants with increased systemic resistance to a wide range of pathogens.  
 CC COI1 and LSI1 proteins may be used to identify inhibitors of these  
 CC proteins, which may be useful as herbicides  
 CC  
 XX

SQ Sequence 2288 BP; 560 A; 510 C; 546 G; 672 T; 0 U; 0 Other;

Query Match		100.0%;	Score 2288;	DB 3;	Length 2288;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 2288;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GCACGAGGCCACACGTTACACAGGCGACTATGGTTGCGGAAACAAATCCGGATGGGAAA	60		
DB	1	GCACGAGGCCACACGTTACACAGGCGACTATGGTTGCGGAAACAAATCCGGATGGGAAA	60		
QY	61	GGGTGTATGTAGTCTGTTCTAGGATGAATATTTGATTAACAGAACGGGGTTTGAAGCAGT	120		
DB	61	GGGTGTATGTAGTCTGTTCTAGGATGAATATTTGATTAACAGAACGGGGTTTGAAGCAGT	120		
QY	121	GACGTGTTACATCAGTACATCATCATCATCATCATCATCATCATCATCATCATCATCAT	180		
DB	121	GACGTGTTACATCAGTACATCATCATCATCATCATCATCATCATCATCATCATCATCAT	180		
QY	181	AGTTTGTGTTTACACCCATCTGTTGGCCCTACCAATTCCTCACTCATCATCCCAT	240		
DB	181	AGTTTGTGTTTACACCCATCTGTTGGCCCTACCAATTCCTCACTCATCATCCCAT	240		
QY	241	AACCCATTCCCTTTTGCCACTTGAACCAAAACCTCTGCACCTTTTCTTTTCACTCTCAG	300		
DB	241	AACCCATTCCCTTTTGCCACTTGAACCAAAACCTCTGCACCTTTTCTTTTCACTCTCAG	300		
QY	301	TCTCCGATCCAATATGACGGAGAACGGAACGTCGGGAAGACACGTCGTGGT	360		
DB	301	TCTCCGATCCAATATGACGGAGAACGGAACGTCGGGAAGACACGTCGTGGT	360		
QY	361	CTCGACTGGTTCATCCCTTACATCGACGACCCCAAGGACCGCGCTTCCAGGT	420		
DB	361	CTCGACTGGTTCATCCCTTACATCGACGACCCCAAGGACCGCGCTTCCAGGT	420		
QY	421	GTGTGACGCTGTGACGAGTCTGCTCGCTCACCCGCAAGCAGTCACCATCGCGCTCTG	480		
DB	421	GTGTGACGCTGTGACGAGTCTGCTCGCTCACCCGCAAGCAGTCACCATCGCGCTCTG	480		
QY	481	CTACACCAACACCCCGGCTCGCTCGCGCGCGCTTCCGCAAGCAGTCACCATCGCGCTCTG	540		
DB	481	CTACACCAACACCCCGGCTCGCTCGCGCGCGCTTCCGCAAGCAGTCACCATCGCGCTCTG	540		
QY	541	CAAGGCGAAGCCCGAGCGCAATGTTCACTTGATACCCGAGAGTTGGGCGGACACGT	600		
DB	541	CAAGGCGAAGCCCGAGCGCAATGTTCACTTGATACCCGAGAGTTGGGCGGACACGT	600		
QY	601	CATCTCGCTGGGTCAAAAGAGATTTCTCAGTACTTTCGATTCGCTCAAGAGCCTCCACTTCG	660		
DB	601	CATCTCGCTGGGTCAAAAGAGATTTCTCAGTACTTTCGATTCGCTCAAGAGCCTCCACTTCG	660		
QY	661	CCGATGATTTGTCAGAGTTCCGATCTTCAAGATCTCGCTCGCTGACCGCGGTCACTGCT	720		
DB	661	CCGATGATTTGTCAGAGTTCCGATCTTCAAGATCTCGCTCGCTGACCGCGGTCACTGCT	720		
QY	721	TCAGCTCTCAAGCTTGACAAAGTCTCCGGTTTCCACCCGATGCTCTTTTCCATATCGG	780		
DB	721	TCAGCTCTCAAGCTTGACAAAGTCTCCGGTTTCCACCCGATGCTCTTTTCCATATCGG	780		
QY	781	TCGCTTTTGAAGAGTTTAAGAGTCTTGTGTTTGGAGGAAGCTCAATTTTGAAGGA	840		
DB	781	TCGCTTTTGAAGAGTTTAAGAGTCTTGTGTTTGGAGGAAGCTCAATTTTGAAGGA	840		
QY	841	CGGAGAAATGCGTACACGAGCTTGTGTTGATATATACAGTCTCTTGAGACTCTCAATTTT	900		
DB	841	CGGAGAAATGCGTACACGAGCTTGTGTTGATATATACAGTCTCTTGAGACTCTCAATTTT	900		

QY	901	CTTGACAGACATTTGCTGTTGTGAAGATTGAGGACTTTGAACTTTTAGCTAAAAATTGCCC	960
DB	901	CTTGACAGACATTTGCTGTTGTGAAGATTGAGGACTTTGAACTTTTAGCTAAAAATTGCCC	960
QY	961	CAACTTAGTGTCTGTGAACACTTACTGCTGTAATCTGTAATCTGTAATCTGTAATCTGTAAT	1020
DB	961	CAACTTAGTGTCTGTGAACACTTACTGCTGTAATCTGTAATCTGTAATCTGTAATCTGTAAT	1020
QY	1021	GCATGCTCTGCGCTGGAAGAGTTTGTGGAGGCACTTACAAACGAGGAAACGAGAAAGATA	1080
DB	1021	GCATGCTCTGCGCTGGAAGAGTTTGTGGAGGCACTTACAAACGAGGAAACGAGAAAGATA	1080
QY	1081	CTCTGCTATATCATTTACAGCAAAAGTTATGTCGATGGGTTTAAATATATATGGAAGAA	1140
DB	1081	CTCTGCTATATCATTTACAGCAAAAGTTATGTCGATGGGTTTAAATATATATGGAAGAA	1140
QY	1141	TGAGTGGCCCATTTGTTTCATGTTTGAGCGCTACTAAAAAATTTGGATCTCTCTATGC	1200
DB	1141	TGAGTGGCCCATTTGTTTCATGTTTGAGCGCTACTAAAAAATTTGGATCTCTCTATGC	1200
QY	1201	AATGCTAGACGAGGAGTCAATGTTTAAATCCAAAGGTCTCCAAATCTGGAAGTCT	1260
DB	1201	AATGCTAGACGAGGAGTCAATGTTTAAATCCAAAGGTCTCCAAATCTGGAAGTCT	1260
QY	1261	TGAGCAAGGAATGTAAATTTGAGATAGAGGTTAGAGTTCTTGGTCTGTTGTAAGAG	1320
DB	1261	TGAGCAAGGAATGTAAATTTGAGATAGAGGTTAGAGTTCTTGGTCTGTTGTAAGAG	1320
QY	1321	GCTAAAAAGGCTTAGGATTGAAAGGGCGATGATCAAGGAATGGAAGGATGAAGAAGG	1380
DB	1321	GCTAAAAAGGCTTAGGATTGAAAGGGCGATGATCAAGGAATGGAAGGATGAAGAAGG	1380
QY	1381	TACTGTGCTCCATAGAGGCTTAATAGCTTGTGTCAGGGCTGTTCAGAGCTTGAATACAT	1440
DB	1381	TACTGTGCTCCATAGAGGCTTAATAGCTTGTGTCAGGGCTGTTCAGAGCTTGAATACAT	1440
QY	1441	GGCTGTTTATGTTCTGATATTAATAATGATCTCTGGAACATATTTGGAACCTCACTGAA	1500
DB	1441	GGCTGTTTATGTTCTGATATTAATAATGATCTCTGGAACATATTTGGAACCTCACTGAA	1500
QY	1501	GAACTCTGTGATTTTTCGCTTGTGTTGCTTGCACCATGAAAGAGAGATACTGATTGCC	1560
DB	1501	GAACTCTGTGATTTTTCGCTTGTGTTGCTTGCACCATGAAAGAGAGATACTGATTGCC	1560
QY	1561	ACTTGACAAATGAGGCTGAGGCTCTACTGAGGGCTGTGACAGAGGATTTGCTCT	1620
DB	1561	ACTTGACAAATGAGGCTGAGGCTCTACTGAGGGCTGTGACAGAGGATTTGCTCT	1620
QY	1621	ATATCTCAGGCGTGGGGGTTGATGATGAGGCTTGTGTTTACATTTGGAACAATACAGTCC	1680
DB	1621	ATATCTCAGGCGTGGGGGTTGATGATGAGGCTTGTGTTTACATTTGGAACAATACAGTCC	1680
QY	1681	AAATGTGAGATGATGCTGTTGTTTATGCTGGGGAGTCTGATGCAAGGCTTTTGGAGTT	1740
DB	1681	AAATGTGAGATGATGCTGTTGTTTATGCTGGGGAGTCTGATGCAAGGCTTTTGGAGTT	1740
QY	1741	CGCTAAGGGGTGCTCTAGTCTTTCAGAACTTGAATGAGAGGTTTATTTTTCAGTGA	1800
DB	1741	CGCTAAGGGGTGCTCTAGTCTTTCAGAACTTGAATGAGAGGTTTATTTTTCAGTGA	1800
QY	1801	AGTGACATCTGCTGGCTGCAACAAATTTGATCTCTCTAGTACTTCTGGGTGCAAGG	1860
DB	1801	AGTGACATCTGCTGGCTGCAACAAATTTGATCTCTCTAGTACTTCTGGGTGCAAGG	1860
QY	1861	TTATGTTGATCTTCCATCTGGAAGTCTTCTTTTGGTAATGGCTCCGACCTTTTGGAAAT	1920
DB	1861	TTATGTTGATCTTCCATCTGGAAGTCTTCTTTTGGTAATGGCTCCGACCTTTTGGAAAT	1920
QY	1921	TGAGTTGATTTCTTCTAGAAAGTGGCTACGAATACCAATCCAGATGAGACTGTAGTTGT	1980
DB	1921	TGAGTTGATTTCTTCTAGAAAGTGGCTACGAATACCAATCCAGATGAGACTGTAGTTGT	1980









DR	P-PSDB; AAB23456.	
XX	New nucleic acid sequences encoding new disease resistance factors,	
PT	useful for producing plants with increased resistance to pathogens and to	
PT	screen for herbicides.	
XX	Claim 2; Page 55; 74pp; English.	
XX	Expressed sequence tags (ESTs) from rice, wheat, soybean and corn cDNA	
CC	libraries were found to be similar to cDNA encoding the Arabidopsis	
CC	thaliana COI1 protein and the Zea mays L1S1 protein. The present sequence	
CC	is cDNA from the corn clone p0128.cpici34r:fls which is homologous to	
CC	COI1. The COI1 and L1S1 proteins are involved in disease resistance. The	
CC	COI1 and L1S1 DNAs of the invention may be used to alter the expression	
CC	of COI1 and L1S1 protein in cells, particularly to produce transgenic	
CC	plants with increased systemic resistance to a wide range of pathogens.	
CC	COI1 and L1S1 proteins may be used to identify inhibitors of these	
XX	proteins, which may be useful as herbicides	
SQ	Sequence 1728 BP; 496 A; 341 C; 400 G; 491 T; 0 U; 0 Other;	
	Query Match 16.3%; Score 372.4; DB 3; Length 1728;	
	Best Local Similarity 61.0%; Pred. No. 6.5e-94;	
	Matches 702; Conservative 0; Mismatches 421; Indels 27; Gaps 5;	
QY	807 TGTCTTTGAGGAAAGCTCAATCTTTGAGAAG--ACGGAGAAATGGCTACACGAGCTTG 863	
DB		
QY	864 CTTTGAATATACAGTCTTTGAGACTCTCAATTTTACTTTGACAGACATGCTGTGTGA 923	
DB		
QY	924 AGATTGAGACCTTGAACTTTTGTAGCTAAATAATGCCCCACTAGTCTGTGAACTTA 983	
DB		
QY	154 CA---GAGTATCTTACCCCTCTTGTGCGCAACTGTCAACGATTGAAACCTCTGAAGATTA 210	
DB		
QY	984 CTGACTGTGAATACAGTCTTGTGAACTCTTGAAGCTCTTGAAGCATGCTCTGGCTGGAAGAT 1043	
DB		
QY	211 GTGAATGTTTCATGCCGATCTGCTGATGTTTTCGAACTGCAACACACTACAGAGT 270	
DB		
QY	1044 TTTGTGGAGCACCTCAACAGAGAACCCAGAAAGATCTCTGTCTATA----- 1090	
DB		
QY	271 TCGCTGGTGTCTCTTGAAGAGCAGGCTCAACCTGTGGCAAGTAGAATATGAGAACT 330	
DB		
QY	1091 --TCATTACAGCAAGATTTATGCTGATGCTGTTTACATATATTGGAAGATGAGTTGC 1148	
DB		
QY	331 ACTATTTCTCTCTTCACTGCAACGCTTGTGCTCTACATGGGAACAAATGATATGC 390	
DB		
QY	1149 CCATTGTGTTTCATGTTTGCAGCGCTACTAAAAAATTTGGATCTCCTCTATGCAATGCTAG 1208	
DB		
QY	391 AATATCTGTTTCCATATGCTACTGCACTTGAAGATTTAGACCTTCAGTTTACATTCCTTT 450	
DB		
QY	1209 ACACGGAGGATCATTTGATGTTTAAATCCAAAGGTGTCAAATCTGGAAGTCTTTCAGACAA 1268	
DB		
QY	451 CCACAGAGGATCATTTGTGATGATGTTCAACGCTGCTCCAATCTGGAACCTTTAGAGTGA 510	
DB		
QY	1269 GGBATGTTAATTTGAGATAGAGGTTTACAGTCTTCTGCTGTTTGTGAAGAGGCTAAAA 1328	
DB		
QY	511 GGGATGTCATAGGGGATCGTGGATCAAGTTTGTTCACAGACCTTCGAAGAAATTCATTA 570	
DB		
QY	1329 GGCTTAGATTTGAAGGGGGGATGATGATCAAGCA---ATGGAGGATGAAGAGGTTACTG 1385	
DB		
QY	571 GGCTCAGATGAGAGAGGAGATGATGATCAAGGAGGTTTGAAGATGAACAGGTTAGGA 630	
DB		
QY	1386 TGTCCCATAGAGGCTAATAGCCTTTGTACAGGCTGTTTCAGACTTGAATACATGCTG 1445	
DB		
QY	631 TTTACAGAGTGGGTTGATGCTATAGCCCAAGGCTGCCCTGAGTTGACATCTACTGGCGGA 690	
DB		
QY	1446 TTTATGCTGATATACAAATGCTCTCTGGAACATATTGGAACCTCACTTGAAGACC 1505	
DB		
QY	691 TACATGTATCAGACATTAACAAATCGCTTTTAGAGGCGAGTTTGGTACATGACAGCAAAATC 750	
DB		

QY	1506 TCTGTGATTTTCGGCTTGTGTTGCTTGCACATGAGAGAAATCACTGATTTGCCACTTG 1565	
DB		
QY	751 TTAATGACTTCGGCTTGTCTCTTGTATAGAGAGCACATATAACCGAATTCGCCACTGG 810	
DB		
QY	1566 ACAATGGGCTGAGGCTCTACTAGGGGCTGTGCAAGCTGAGGAGATTTGCTCTATATC 1625	
DB		
QY	811 ACAATGGGCTTCGTGCTTTGCTTTAGAGGTTGCACCAACTACGGAGGTTTGCATTTTATG 870	
DB		
QY	1626 TCAGGCTGGGGGTTGACTGATGTAGGCTTGGTTTACATTTGGACAATACAGTCCAAATG 1685	
DB		
QY	871 TGAGACCTGGGGCCCTATCTGATTTGCTTGGCTATGTTGGAGAATTTAGTAAGATGA 930	
DB		
QY	1686 TGAGATGATGCTGCTTGGTTTATGTGGGGAGTCTGATGAGGGCTTTTGGAGTTGCTTA 1745	
DB		
QY	931 TTCGTTATATGTTGCTTGGTTAATGTTGGTGAATCTGATAATGAATCATACAAATTTATCAA 990	
DB		
QY	1746 AGGGGTGCTTCTAGTCTTCAGAACTTGAATGAGAGGGTGTATTATTTTCAGTGAACGTG 1805	
DB		
QY	991 AAGGCTGCCCAAGCTTGCACAAACTTGGAGGTGAGGGTGTGCTCTTT---AGTGAGCATG 1047	
DB		
QY	1806 CACTTGTGCTGCTGCAACAAATTTACTTCTTTAGGTACTTGTGGTGCAGGTTTATG 1865	
DB		
QY	1048 CTTTACGCTTGGCTGCACTACAGCTTAAGTCACTGAGGTATCTGTGGGTACCAAGGATTCA 1107	
DB		
QY	1866 GTGATCTCCATCTCGAGCTGTATCTTTTGGTATGCTGACCCCTTTTGGACATTTGAGT 1925	
DB		
QY	1108 GGTCACTCTCCAACCTGAACTGATATTATTTGGCAATGTTAGCCCTTCTGGAACATTTGAGT 1167	
DB		
QY	1926 TGATTTCCCTTC 1935	
DB		
QY	1168 ATATTGTCTCC 1177	
DB		
XX	RESULT 6	
XX	AA02372	
ID	AA02372 standard; cDNA; 844 BP.	
XX	AC	
XX	AA02372;	
DT	30-OCT-2001 (first entry)	
XX	COI1 contig of r12.pk0027.h4, r10n.pk099.p14 and r10n.pk0047.c5.	
DE	COI1; rice; herbicide; disease-resistance factor; de.	
KW	Oryza sativa.	
OS	Location/Qualifiers	
PH	9..715	
FT	/*tag= a	
FT	/product= "COI1"	
FT	/partial	
FT	/tranel except= (pos:352..354,aa:Xaa)	
FT	/tranel "Xaa= unknown"	
FT	/tranel except= (pos:663..664,aa:Arg)	
FT	/note= "this codon has an apparent 1 nucleotide deletion	
FT	which alters the reading frame"	
XX	WO200068406-A2.	
PN	16-NOV-2000.	
XX	03-MAY-2000; 2000WO-US011956.	
XX	07-MAY-1999; 99US-0133041P.	
XX	(DUPO ) DU PONT DE NEMOURS & CO E I.	
XX	Caimi PG, Famodu OO, Lee J, Miao G, Maxwell CA;	
XX	WPI; 2000-687649/67.	
DR	P-PSDB; AAP01920.	
XX		

PT New nucleic acid sequences encoding new disease resistance factors,  
 PT useful for producing plants with increased resistance to pathogens and to  
 PT screen for herbicides.

XX Claim 2; Page 44; 74pp; English.

XX An isolated polynucleotide (I) comprising a sequence (Ia) encoding a  
 CC polypeptide (II) of at least 60 amino acids (aa) from any of 18 specified  
 CC sequences, given in the specification, or the complements of (Ia), is  
 CC new. Herbicide. (I) encodes disease-resistance factor polypeptides (II)  
 CC and can be used, in sense or antisense orientation, to alter the  
 CC expression of (II) in cells, particularly to produce transgenic plants  
 CC with increased systemic resistance to a wide range of pathogens. (I) or  
 CC their fragments are also used for gene mapping (e.g. for plant breeding)  
 CC and to identify loss of function mutants. Polypeptides (II) encoded by  
 CC (I) are used to raise specific antibodies. For detection of (II) and to  
 CC design and/or identify specific inhibitors of (II), potentially useful as  
 CC herbicides. This sequence encodes the rice COI1 protein created from a  
 CC contig of rlr2.pk0027.h4, rlon.pk0099.p14 and rlon.pk0047.c5 as described  
 CC in the method of the invention

XX Sequence 844 BP; 223 A; 171 C; 233 G; 215 T; 0 U; 2 Other;

Query Match 13.0%; Score 296.6; DB 3; Length 844;  
 Best Local Similarity 65.5%; Pred. No. 1e-72;  
 Matches 464; Conservative 0; Mismatches 240; Indels 4; Gaps 2;

QY 1245 CAATCTGGAAGTCTTGAGACAGGATGTAATTTGGAGATAGAGGGTTAGAGGTTCTTG 1304  
 DB 7 CCACCTACTAGTTCTTCGGTGAGGAATGTAATTTGGAGATAGAGGTTAGGGTTGTTG 66  
 QY 1305 GTCGTTGTTGTAAGAGGCTAAAGGCTTAGGATTGAAGGGCGGATGATGATCAAGAA 1364  
 DB 67 CAGACACATGCAAGAAGCTCAAGAGCTCAGAGTTCAGGAGGAGATGATGATCCAGTT 126  
 QY 1365 TGGAGGATGAAGAGGTAAGTCTGTCCTAGAGGCTTAATAGCCTTGTCACAGGCTGTT 1424  
 DB 127 TGCAGAGAAACAGAGGAGGATCTCAAGTGGGTTGACAACTGTAGCCGTAGATGCC 186  
 QY 1425 CAGAGCTTGATACATGCTGTTTATGTCGTGATTTACAAATGCTCTGGAACATA 1484  
 DB 187 GTGAACCTGGAATACATAGCTGCTTAAGTCTGATATCAAAATGGGCGCTGAGTCTA 246  
 QY 1485 TTGGAACCTCATTGGAAGAACCTCTGTGATTTTCGCTTGTGTTGATGACCATGAAGAGA 1544  
 DB 247 TTGGGACCTTCGCAAAATCTTTCGCACTTCGCTTGTCTCTCGATGAGAGAGA 306  
 QY 1545 AGATAACTGATTTGCCACTTGACATGGGCTGAGGCTCTACTGAGGCGCTGTGACAGC 1604  
 DB 307 GGATAACAGATTTGCCCTTAGACAAATGTTGTCGCTGCACTGTGANGGCTGCACGAAC 366  
 QY 1605 TCAGGAGATTTGCTCTATATCTCAGCGTGGGCTGCTGACTGATGAGGCTTGTGTACA 1664  
 DB 367 TTCGAGGTTTGTCTTATCTTGACACAGGGGACTTCAGATACAGGCTTGTGCTATA 426  
 QY 1665 TTGGCAATFACAGTCCAAATGTGAGATGATGCTGCTGTTTATGTTGGGGAGTCTGTG 1724  
 DB 427 TTGGCAGTACAGTGGAAATATCCAATACATGCTTCGGGTAATGTTGGGGAACAGATG 486  
 QY 1725 CAGGGCTTTTGAGTTGCTGTAAGGGGTGTCCTAGTCTTCAGAACTTGAATGAGAGGT 1784  
 DB 487 ATGGTCTGATCCGGTTTGTGATTTGGGGGTGTGAGAACCTGCGAAGCTTGA--GCTAAGA 543  
 QY 1785 GTTATATTTTCAGTGAACGTGCACTTGTGCTGCAACACAAATGACTTCTCTTAGT 1844  
 DB 544 GTTGTGCTTCACTGAGCAAGCTTTAGCCCGGCTATACGGATATGCTTCCCTGAGAT 603  
 QY 1845 ACTTGTGGGTGCAAGGTTATGTTGATCTCCATCTTGTGGAAGTATCTTTTGGTAAGGCTC 1904  
 DB 604 ACCTGTGGGTACAGGGCTCAAGGCTTCTAAGACTGCTCAGCATCTCATGCTCATGCC-A 662  
 QY 1905 GACCTTTTGGACATGATGTTGATTTCTTCTAGAAAGGTGGCTACGA 1952

DB 663 GGCCCTCTTGGACATAGAGTTTACACCTCCACAGAGACTGGTCAAGA 710

RESULT 7

AAA95054

ID AAA95054 standard; cDNA; 844 BP.

XX AAA95054;

DT 12-FEB-2001 (first entry)

XX cDNA from a rice contig homologous to COI1.

DE Corn; rice; soybean; wheat; ATHCOR1; LLS1; disease resistance;  
 KW A. thaliana coronatine-induced; COI1; herbicide; EST;  
 KW expressed sequence tag; contig; ss.

OS Oryza sativa.

XX Location/Qualifiers

FT Key 9..715

FT CDS /\*tag= a

FT /product= "COI1"

FT /transl\_except= (pos:663..664,aa:Arg)

FT /partial

XX WO200068406-A2.

XX 16-NOV-2000.

XX 03-MAY-2000; 2000WO-US011956.

XX 07-MAY-1999; 99US-0133041P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Calmi PG, Famodu OO, Lee J, Miao G, Maxwell CA;

XX WPI; 2000-687649/67.

DR P-PSDB; AAB23449.

XX New nucleic acid sequences encoding new disease resistance factors,  
 PT useful for producing plants with increased resistance to pathogens and to  
 PT screen for herbicides.

XX Claim 2; Page 44; 74pp; English.

XX Expressed sequence tags (ESTs) from rice, wheat, soybean and corn cDNA  
 CC libraries were found to be similar to cDNA encoding the Arabidopsis  
 CC thaliana COI1 protein and the Zea mays LLS1 protein. The present sequence  
 CC is cDNA from a rice contig homologous to COI1. The COI1 and LLS1 proteins  
 CC are involved in disease resistance. The COI1 and LLS1 DNAs of the  
 CC invention may be used to alter the expression of COI1 and LLS1 protein in  
 CC cells, particularly to produce transgenic plants with increased systemic  
 CC resistance to a wide range of pathogens. COI1 and LLS1 proteins may be  
 CC used to identify inhibitors of these proteins, which may be useful as  
 CC herbicides

XX Sequence 844 BP; 223 A; 171 C; 233 G; 215 T; 0 U; 2 Other;

Query Match 13.0%; Score 296.6; DB 3; Length 844;

Best Local Similarity 65.5%; Pred. No. 1e-72;

Matches 464; Conservative 0; Mismatches 240; Indels 4; Gaps 2;

QY 1245 CAAATCTGGAAGTCTTGAGACAGGATGTAATTTGGAGATAGAGGTTAGAGGTTCTTG 1304

DB 7 CCACCTACTAGTTCTTCGGTGAGGAATGTAATTTGGAGATAGAGGTTAGGGTTGTTG 66

QY 1305 GTCGTTGTTGTAAGAGGCTTAAAGGCTTAGGATTGAAGGGCGGATGATGATCAAGAA 1364

DB 67 CAGACACATGCAAGAAGCTTACAAAGACTCAGAGTTGACGAGAGATGATGATCCAGTT 126

QY 1365 TGGAGGATGAAGAAGGTACTGTGTGCCATAGAGGGCTAATAGCCTTGTTCACAGGGCTGTT 1424

```

Db 127 TCGAAGAGAACAGAGGAGGTCTCTCAAGTCGGGTTGACAACTGTAGCCGTAGATGCC 186
QY 1425 CAGAGCTTGAATACATGCTGTTTATGCTGTGATTAACAAATCATCTCTGGAACATA 1484
Db 187 GTGAATCGAATACATGCTGCTTAAGTCTGATATACAAATGGGCCCTTGGAGTCTA 246
QY 1485 TTGGAACACTCATTGAAGAACCTCTCTGATTTTTCGGCTTGTGTTGCTTGAACATGAAGAGA 1544
Db 247 TTGGGACTTCTGCAAAATCTTTGGGACTTCCGCTTGTCTACTCGATAGAGAAGA 306
QY 1545 AGATTAACATGTTGCCACTTGAATAGGCTGAGGGCTCTACTAGAGGGCTGTGACAAGC 1604
Db 307 GGATAACAGATTTGGCCTTAGACAAATGGTGTCCGCTGCACTGCTGANGGGCTGCACGAAAC 366
QY 1605 TCAGAGAGATTTGCTCTATATCTCAGGCGTGGGGGCTTGAATGATGTAGGCCCTTGGTTACA 1664
Db 367 TTCGAGAGTTTCTCTATATCTTGAACACGAGGGGACTTTCAGATACAGGCCCTTGGCTATA 426
QY 1665 TTGGACAATACAGTCCAAATGTGAGATGATGCTGCTTGGTTTATGTTGGGGAGTCTGATG 1724
Db 427 TTGGACAGTACAGTGGAAATATATCAATACATGCTTCTGGGTAAATGTTGGGAAACAGATG 486
QY 1725 CAGGCGTTTTCAGTTCGCTAGAGGGGTCTCTAGTCTTCAGAACTTGAATGAGAGGT 1784
Db 487 ATGGTCTGATCCGGTTTGCATTTGGGGGTGTGAAACCTTCGGGAGGCTTGA---GCTAAGGA 543
QY 1785 GTTTATTTTTCAGTGAACGTGCACTGTCTGTGGCTGCAACACAAATGTACTTCTCTAGGT 1844
Db 544 GTTGTGCTTTCAGTGAAGCTTTAGCCGCGCTATACGGAGTATGCTTCTCTCTAGAT 603
QY 1845 ACTGTGGGTGAAGTTAGTGTATCTCCATCTGAGAGTGAATCTTTTGGTTAATGGCTC 1904
Db 604 ACGTGTGGGTACAGGGCTACAAAGGCTTCTAAGACTGGTCAAGATCTCATGCTCATGGC-A 662
QY 1905 GACCCCTTTTGGACATTTGATTTCTCTCTAGAAAGTGGCTAGGA 1952
Db 663 GGCCCTTCTGGAACATAGAGTTTACACCTCCAGAGAGCTGTGTACGA 710

```

## RESULT 8

ID AAA95066 standard; cDNA; 1074 BP.

AC AAA95066;

DT 12-FEB-2001 (first entry)

DE cDNA from the wheat clone wreln.pk0122.d3.fis homologous to COIL.

XX Corn; rice; soybean; wheat; ATHCOR1; LLS1; disease resistance;  
 KW A. thaliana coronatine-induced; COI1; herbicide; EST;  
 KW expressed sequence tag; wreln.pk0122.d3.fis; ss.

OS *Triticum aestivum*.

XX Key Location/Qualifiers  
 FH 2..667  
 CDS /\*tag= a  
 FT /product= "COIL"  
 FT /partial

XX W0200068406-A2.

PN 16-NOV-2000.

XX 03-MAY-2000; 2000WO-US011956.

XX 07-MAY-1999; 99US-0133041P.

XX (DUPO ) DU PONT DE NEMOURS &amp; CO E I.

XX Caiji PG, Famodu OO, Lee J, Miao G, Maxwell CA;

XX WPI; 2000-687649/67.  
 DR P-PSDB; AAB23461.

XX New nucleic acid sequences encoding new disease resistance factors,  
 PT useful for producing plants with increased resistance to pathogens and to  
 PT screen for herbicides.

Claim 2; Page 66; 74pp; English.

XX Expressed sequence tags (ESTs) from rice, wheat, soybean and corn cDNA  
 CC libraries were found to be similar to cDNA encoding the Arabidopsis  
 CC thaliana COIL protein and the Zea mays LLS1 protein. The present sequence  
 CC is cDNA from the wheat clone wreln.pk0122.d3.fis which is homologous to  
 CC COIL. The COIL and LLS1 proteins are involved in disease resistance. The  
 CC COIL and LLS1 DNAs of the invention may be used to alter the expression  
 CC of COIL and LLS1 protein in cells, particularly to produce transgenic  
 CC plants with increased systemic resistance to a wide range of pathogens.  
 CC COIL and LLS1 proteins may be used to identify inhibitors of these  
 CC proteins, which may be useful as herbicides

Sequence 1074 BP; 282 A; 232 C; 261 G; 299 T; 0 U; 0 Other;

Query Match 10.2%; Score 233.2; DB 3; Length 1074;

Best Local Similarity 65.7%; Pred. No. 9e-55;

Matches 356; Conservative 0; Mismatches 183; Indels 3; Gaps 1;

QY 1396 AGGGCTAATAGCCCTTGTTCACAGGGCTGTTCAGAGCTTGAATACATATGCTTTATGTGTC 1455

Db 10 AGGATTGATGGCTGTAGCTGAAGGCTGCTCTGATTGGAGTACTGGGCAATGATGTGTC 69

QY 1456 TGATATTCAATGCAATCTCTGGAACATATTGGAATCTCACTTGAAGAACTCTGTGATT 1515

Db 70 TGACATTCAATGCAATCTCTGAGGCTATTGGGCAATTCAGCAAAACCTGAAAGATT 129

QY 1516 TCGCCTTCTGTGTTGCTTGCATGAGAGAAATACTGATTGCCACTTGAACAATGGGGT 1575

Db 130 CGACTTGTCTGCTTGTATAGAGAGTGCATATACTGAACTGCCCTTGCACACGGGT 189

QY 1576 GAGGCTCTACTGAGGGCTGTGACAACTGAGAGATTTGCTCTATATCTCAGGGGTGG 1635

Db 190 TCGGGCTTTGCTGAGAGGTTGCACCAACTCCGAGGTTTGCATTTTATGTGAGACCTGG 249

QY 1636 CGGTTGACTGATGAGGCTTGTGTTACATTTGGAACAATACAGTCCAAATGTGAGATGAT 1695

Db 250 AGCTCTATCAGATATTTGGCTTTCTTATGTTGGCGAATTTAGCAAGACCGTCCGTACAT 309

QY 1696 GCTGCTTGGTTATGTGGGGAGTCTGTATGTCAGGGGCTTTTGGAGTTTCGTAAGGGGTGTC 1755

Db 310 GTTGTCTGGAAATCGCGGGGGTCTGTATGATGACTGTGGCATTTGCACGAGGATGCC 369

QY 1756 TAGTCTTCAGAAAATTGAAATGAGAGGGTGTATTATTTTCACTGAACTGCACTTGTCTGT 1815

Db 370 AAGCTTGCAGAAATTTGGAGCTAA---GGAGTTGCTGCTTTAGTGAACGTCATTGGCAGT 426

QY 1816 GGCTGCAACACAAATTTGACTTCTCTTAGGTACTTGTGGGTGCAAGGTTATGTTGTTATCTCC 1875

Db 427 TGCAGCCTTACAGCTGAAGTCACTCAGATATCTTTGGGTGCGAGGATACAGGCATCTCC 486

QY 1876 ATCTGGAGCTGATCTTTTGGTAAATGGCTCGACCCCTTTTGGAAACATTTGATTCCTTC 1935

Db 487 TACTGGCAGCGATCTCATGGCAATGGTACGCCCTTCTGGAACATTTAGTTTATTGGCACC 546

QY 1936 TA 1937

Db 547 AA 548

## RESULT 9

ABQ65584

ID ABQ65584 standard; DNA; 719 BP.

XX ABQ65584;

AC ABQ65584;



1536 CGGGTTGACTGATGTAGGCCCTGGTTACATTGGAC - AATACAGTCCAAATGTGAGATGGA 1694







EH Key Location/Qualifiers  
FT CDS 8...589  
FT /\*tag= a  
FT /product= "COI1"  
FT /partial  
XX WO200068406-A2.  
XX 16-NOV-2000.  
XX 03-MAY-2000; 2000WO-US011956.  
XX 07-MAY-1999; 99US-0133041P.  
XX (DUPO ) DU PONT DE NEMOURS & CO E. I.  
XX Caiimi PG, Farnodu OO, Lee J, Miao G, Maxwell CA;  
XX WPI; 2000-687649/67.  
XX P-PSDB; AAB23448.  
XX New nucleic acid sequences encoding new disease resistance factors,  
PT useful for producing plants with increased resistance to pathogens and to  
PT screen for herbicides.  
XX Claim 2; Page 42-43; 74pp; English.  
XX Expressed sequence tags (ESTs) from rice, wheat, soybean and corn cDNA  
CC libraries were found to be similar to cDNA encoding the Arabidopsis  
CC thaliana COI1 protein and the Zea mays LLS1 protein. The present sequence  
CC is cDNA from the corn clone p0128.cpic134r which is homologous to COI1.  
CC The COI1 and LLS1 proteins are involved in disease resistance. The COI1  
CC and LLS1 DNAs of the invention may be used to alter the expression of  
CC COI1 and LLS1 protein in cells, particularly to produce transgenic plants  
CC with increased systemic resistance to a wide range of pathogens. COI1 and  
CC LLS1 proteins may be used to identify inhibitors of these proteins, which  
CC may be useful as herbicides  
XX  
SQ Sequence 701 BP; 193 A; 140 C; 179 G; 180 T; 0 U; 9 Other;  
Query Match 5.6%; Score 129; DB 3; Length 701;  
Best Local Similarity 55.5%; Pred. No. 1.7e-25;  
Matches 318; Conservative 0; Mismatches 234; Indels 21; Gaps 3;  
QY 807 TGTITTTGGAGGAAGCTCAATCTTTCAGRAGG---ACGGAGANTGGCTACAGAGCTTG 863  
DB 24 TTTTCTCGGAAGAAAGCAACATTCATGAGAAAGAAATGATGAGTGGATCCGTGAGCTTG 83  
QY 864 CTTTGAATATACAGTTCTTTCAGACTCTCAATTTTCTTTCACAGACATTCCTTCTTGA 923  
DB 84 CTACGAGCAATTCGTCTTTCAGACTGACTGANTTTCTTTCACAGATCTCAGGSCATCCC 143  
QY 924 AGATTGAGGACCTTTGAACTTTTACATAAAATGCCCAACTTAGTGTCTGTGAACTTA 983  
DB 144 CA---GAGTATCTTACCTCTCTGTGGCAACTGTCAACGATTGAAAACTCTGAAGATTA 200  
QY 984 CTGACTGTGAATACCTGGATCTTTCAGACTCTTTTACGATCGCTCTCGCTGGAAGAGT 1043  
DB 201 GTGAATGTTTCATCCCGCATCTGGTCACTTGTTCGGAACCTGCACAACTACAGAGT 260  
QY 1044 TTTGTGAGGACCTTACAAACGAGGAACCCAGAAAGATATCTCTGCTATA----- 1090  
DB 261 TCGCTGGTGGTTCCTTTTGAAGAGCAGGGTCAACCTGTGGCAAGTAGAAATTTATGAGAACT 320  
QY 1091 ---TCATTACAGCAAGATTATGTGANTGGGTTTAACTATATTTGGAAGAAATGAGTTGC 1148  
DB 321 ACTATTTTCTCTCACTGCACCGCTTGAGTTTGTCTCTACATGGGAACAAATGATATGC 380  
QY 1149 CCAATGTGTTTCAATGTTTCAGCCGCTACTTAAATAAATTCGATCTCTCTATCAATGCTAG 1208  
DB 381 AAATACTGTTTCCATATCTACTGCACCTTAAGAAGTTAGACCTTCAGTTTACATTCCTTT 440  
QY 1209 ACACGAGGATCATTTGTTATGTTTAATCCAAAGGTGTCCAAATCTGGAAGTCTCTGAGACAA 1268

DB 441 CCACAGAGGATCATTCNCAGATAGTTCAACGGCTGCTCCAATCTGGAACCTTAGAGTGA 500  
QY 1269 GGAATGTAATTGGAGATAGAGGGTTAGAGGTTCTTGGTCTTGTGTAAGAGGCTAAAAA 1328  
DB 501 GGGATGTCATAGGGGATCGGGGACTACAAGNTGGTGCACAGACCTGCAAGAAATTCATTA 560  
QY 1329 GGCTTAGGATTTGAAAGGGGCGGATGATGATCAAG 1361  
DB 561 GGCTCAGAGTAGAGAGGAGGATGATGATCAG 593

Search completed: April 19, 2004, 18:52:44  
Job time : 880 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 18:30:24 ; Search time 164 Seconds

(without alignments)  
7742.242 Million cell updates/sec

Title: US-10-009-791-21

Perfect score: 2288

Sequence: 1 gcacgaggccacacgttaca.....acattctcaattttgtgaat 2288

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs. 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Minimum DB seq	length: 9
Maximum DB seq	length: 2000000000

post-processing: Minimum Match 0%

FOR-PROCESSING: Minimum Match 0%  
Maximum Match 100%

Maximum March 1998  
Listing first 45 summaries

Database :

```

1: /cgn2_6/prodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/2/ina/PCUTS_COMB.seq: *
6: /cgn2_6/prodata/2/ina/backfiles1.seq: *

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	105.2	4.6	266	4	US-09-313-294A-1866		Sequence 1866, Ap
2	54	2.4	7218	1	US-08-232-463-14		Sequence 14, Appl
C 3	49.6	2.2	7218	1	US-08-232-463-14		Sequence 14, Appl
4	46.8	2.0	735	3	US-09-003-287-7		Sequence 7, Appl
5	44.8	2.0	1020	4	US-09-170-496D-181		Sequence 181, App
6	44.2	1.9	6085	3	US-09-029-603-4		Sequence 4, Appl
C 7	44.2	1.9	34094	4	US-09-292-034-1		Sequence 1, Appl
8	43.4	1.9	1020	4	US-09-170-496D-31		Sequence 31, Appl
9	43.4	1.9	1900	4	US-09-016-434-1484		Sequence 1484, Ap
10	43.4	1.9	1901	1	US-08-153-848-43		Sequence 43, Appl
11	43.4	1.9	1901	3	US-09-299-843A-43		Sequence 43, Appl
12	43.4	1.9	1901	4	US-09-088-337B-43		Sequence 43, Appl
13	43.4	1.9	1901	5	PCR-US93-11153-43		Sequence 43, Appl
14	43.4	1.9	2453	5	PCR-US95-07180-1		Sequence 1, Appl
C 15	43.2	1.9	51952	3	US-08-947-823-1		Sequence 1, Appl
16	42.6	1.9	609	4	US-09-252-991A-15483		Sequence 15483, A
17	42.6	1.9	1083	4	US-09-252-991A-15422		Sequence 15422, A
C 18	42.6	1.9	1461	4	US-09-252-991A-15258		Sequence 15258, A
19	42.6	1.9	1929	4	US-09-252-991A-15533		Sequence 15533, A
20	42.2	1.8	1155	2	US-08-387-942C-7		Sequence 7, Appl
21	42	1.8	1683	4	US-09-252-991A-8337		Sequence 8337, Ap
C 22	42	1.8	1833	4	US-09-252-991A-8371		Sequence 8371, Ap
23	42	1.8	1980	4	US-09-252-991A-8507		Sequence 8507, Ap
24	41.8	1.8	44377	2	US-08-804-227C-7		Sequence 7, Appl
25	41.8	1.8	44377	2	US-08-804-198-1		Sequence 1, Appl
C 26	40.8	1.8	1926	4	US-09-249-585A-2		Sequence 2, Appl
27	40.8	1.8	1926	4	US-09-410-399-3		Sequence 3, Appl

## ALIGNMENTS

## RESULT 1

```

RECORD: 1
US-09-313-294A-1866
; Sequence 1866, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 1866
; LENGTH: 266
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700551651H1
US-09-313-294A-1866

```

Query Match	4.6%;	Score	105.2;	DB	4;	Length	266;
Best Local Similarity	64.7%;	Pred. No.	3.3e-20;				
Matches	172;	Conservative	0;	Mismatches	93;	Indels	1;
						Gaps	1;
QY	1268	AGCAATGTAATTCGACATAGAGGGTTAGAGGTTCTTGTCGTTGTTGTAAGAGCGCTAAAA	1327				
Db	2	AGCAATGTGATTGGAGATAGAGGATTAGAGTTGTTGCCGATACGTCACAGAAGCTCCAA	61				
QY	1328	AGGCTTAGGATTCGAAAGGGCGCATGATGATCAAGGAATGGAGGATGAAGAAGTACTGTG	1387				
Db	62	AGGCTCAGATAGAGCGAGGAGATGATGAAGGAGTGTGCAAGAAG-ACAGGGAGGGGTC	120				
QY	1388	TCCCATAGAGGGCTAATAGCTTGTGCACAGGGCTGTTACAGAGCTTGAATACATGGCTGTT	1447				
Db	121	TCTCAAGTGGGCTTGACGGCTATAGCCGTAGGTTGCCGTGAGCTGGAATATATAGCTGCC	180				
QY	1448	TATGTCTCTGATATTCAAAATGCATCTCTGGACATATTGGAACTCACTTTGAAGAACCCT	1507				
Db	181	TATGTGTCTGATATACCAATGGGCTTGGAAATCTATCGGACATCTTCGCCAAAAACTA	240				
QY	1508	TGTGATTTTCGCCCTTGTTGTTCTTGA	1533				
Db	241	TACGACTTCGGGCTTGTTCTACTTGA	266				

RESULT 2  
 PS-08-212-463-14

RESULT 2  
US-08-232-463-14

[illegible]

	;	Sequence 14, Application US/08232463	
	;	Patent No. 5670367	
	;	GENERAL INFORMATION:	
	;	APPLICANT: DORNER, F.	
	;	APPLICANT: SCHEIFLINGER, F.	
	;	APPLICANT: FALKNER, F. G.	
	;	TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS	
	;	NUMBER OF SEQUENCES: 52	
	;	CORRESPONDENCE ADDRESS:	
	;	ADDRESSEE: Foley & Lardner	
	;	STREET: 1800 Diagonal Road, Suite 500	
	;	CITY: Alexandria	
	;	STATE: VA	
	;	COUNTRY: USA	
	;	ZIP: 22313-0299	
	;	COMPUTER READABLE FORM:	
	;	MEDIUM TYPE: Floppy disk	
	;	COMPUTER: IBM PC compatible	
	;	OPERATING SYSTEM: PC-DOS/MS-DOS	
	;	SOFTWARE: PatentIn Release #1.0, Version #1.25	
	;	CURRENT APPLICATION DATA:	
	;	APPLICATION NUMBER: US/08/232,463	
	;	FILING DATE:	
	;	CLASSIFICATION: 435	
	;	PRIOR APPLICATION DATA:	
	;	APPLICATION NUMBER: US/07/935,313	
	;	FILING DATE:	
	;	APPLICATION NUMBER: EP 91 114 300.6	
	;	FILING DATE: 26-AUG-1991	
	;	ATTORNEY/AGENT INFORMATION:	
	;	NAME: BENT, Stephen A.	
	;	REGISTRATION NUMBER: 29,768	
	;	REFERENCE/DOCKET NUMBER: 30472/114 IMMU	
	;	TELEPHONE: (703)836-9300	
	;	TELEFAX: (703)683-4109	
	;	TELEX: 899149	
	;	INFORMATION FOR SEQ ID NO: 14:	
	;	SEQUENCE CHARACTERISTICS:	
	;	LENGTH: 7218 base pairs	
	;	TYPE: nucleic acid	
	;	STRANDEDNESS: single	
	;	TOPOLOGY: linear	
	;	IMMEDIATE SOURCE:	
	;	CLONE: pTZgpt-F1s	
	;	US-08-232-463-14	
		Query Match 2.4%; Score 54; DB 1; Length 7218;	
		Best Local Similarity 1.6%; Pred. No. 9.4e-05;	
		Matches 6; Conservative 220; Mismatches 140; Indels 0; Gaps 0;	
QY	170	AGCTCGGAAAAGTTTGTTCACACCACCTGTTGGCCCTACCATCTCCTACTC 229	
DB	1057	AGCTTGCGATTT 1116	
QY	230	ATCATCCCCATAACCAATCCCCTTTGCCACTTGAACCAAACCTCGCACCTTTCTT 289	
DB	1117	TT 1176	
QY	290	TTCACTCTCAGTCTCGATCCAAATATGACGGAGAACGGAACGTCGCGAAGACACGTG 349	
DB	1177	TT 1236	
QY	350	GTCGACGTGGTCTCGACTGCCTATCCCTTACATGACGACCCCAAGACCGGACGCC 409	
DB	1237	TT 1296	
QY	410	GTTTCCAGGTGTGTCGACGTGGTAGAGCTGCAGCTCCTCACCAGCAACGTCACCC 469	
DB	1297	TT 1356	
QY	470	ATCGCGCTCTGCTACACCAACCCCGGCTCGCCTCCGCGCGCGCTTCCCGCACTCGAG 529	

  

	;	Sequence 14, Application US/08232463	
	;	Patent No. 5670367	
	;	GENERAL INFORMATION:	
	;	APPLICANT: DORNER, F.	
	;	APPLICANT: SCHEIFLINGER, F.	
	;	APPLICANT: FALKNER, F. G.	
	;	TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS	
	;	NUMBER OF SEQUENCES: 52	
	;	CORRESPONDENCE ADDRESS:	
	;	ADDRESSEE: Foley & Lardner	
	;	STREET: 1800 Diagonal Road, Suite 500	
	;	CITY: Alexandria	
	;	STATE: VA	
	;	COUNTRY: USA	
	;	ZIP: 22313-0299	
	;	COMPUTER READABLE FORM:	
	;	MEDIUM TYPE: Floppy disk	
	;	COMPUTER: IBM PC compatible	
	;	OPERATING SYSTEM: PC-DOS/MS-DOS	
	;	SOFTWARE: PatentIn Release #1.0, Version #1.25	
	;	CURRENT APPLICATION DATA:	
	;	APPLICATION NUMBER: US/08/232,463	
	;	FILING DATE:	
	;	CLASSIFICATION: 435	
	;	PRIOR APPLICATION DATA:	
	;	APPLICATION NUMBER: US/07/935,313	
	;	FILING DATE:	
	;	APPLICATION NUMBER: EP 91 114 300.6	
	;	FILING DATE: 26-AUG-1991	
	;	ATTORNEY/AGENT INFORMATION:	
	;	NAME: BENT, Stephen A.	
	;	REGISTRATION NUMBER: 29,768	
	;	REFERENCE/DOCKET NUMBER: 30472/114 IMMU	
	;	TELEPHONE: (703)836-9300	
	;	TELEFAX: (703)683-4109	
	;	TELEX: 899149	
	;	INFORMATION FOR SEQ ID NO: 14:	
	;	SEQUENCE CHARACTERISTICS:	
	;	LENGTH: 7218 base pairs	
	;	TYPE: nucleic acid	
	;	STRANDEDNESS: single	
	;	TOPOLOGY: linear	
	;	IMMEDIATE SOURCE:	
	;	CLONE: pTZgpt-F1s	
	;	US-08-232-463-14	
		Query Match 2.2%; Score 49.6; DB 1; Length 7218;	
		Best Local Similarity 14.4%; Pred. No. 0.0017;	
		Matches 71; Conservative 202; Mismatches 216; Indels 3; Gaps 1;	
QY	1124	ACATATATGGAAGAATGAGTTCGCCCATCTGTTCATGTTTGCAGCGCTACTTAATAA 1183	
DB	1580	ACATATATCTTAATAATGAAAGTGTGTTGACTAGCGTAGCTCGCTCTAGACGCA 1521	
QY	1184	TTGGATCTCCTCTATGCAATCTGACACGAGAGA---TCATTGTATGTTTAATCCAAGG 1240	
DB	1520	TCTATTTCAGTTTCAAAAAACCGCATGTAGGCATCATCTGTAATTACCTATCTATGCA 1461	
QY	1241	TGTCACAAATCTGGAAGTCTCTTGAG	











```

; APPLICANT: Schweikart, Vicki L.
; TITLE OF INVENTION: Novel Seven Transmembrane Receptors
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11153
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/977,452
; FILING DATE: 17-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Noland, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31794
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1901 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 701..1717
; PCT-US93-11153-43

```

```

Query Match 1.9%; Score 43.4; DB 5; Length 1901;
Best Local Similarity 47.3%; Pred. No. 0.046;
Matches 131; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

QY 344 CGTGTGGTCGACGTGGTCTCGACTGCGTCTATCCCTTACATCGACGACCCCAAGGACCGC 403
Db 1403 CGCATGATCGCCATAGTGTGGCCATCTTCTGGTCTGTTGTCCTTACACGTCAAC 1462

QY 404 GACGCGGTTTCCAGGTTGTGCGAGCTGTGTAAGAGCTCGACTCGCTACCCGCAAGCAC 463
Db 1463 CGCTCCGTCTAGTGTGCACTACCGGACGCCATGGGGCTCTTGGCCACCCAGCGCATC 1522

QY 464 GTCACATCGCGCTGTGTACACACCCCGGCTCGCTCGCGCGGCTTCCCGCAC 523
Db 1523 CTGGCCCTGGCAACCGCATACCTCTCTGCTTCAACAGGCTCAAGGGGCACTCGACCC 1582

QY 524 CTCGAGTCGCTCAAGCTCAAGGGCAAGCCCGAGCGCAATGTTCACCTTGATACCCGAG 583
Db 1583 ATCATGTATTCTTCTGTGCTGAGAAGTTCGCCAGCCCTGTGCAACTTGTCTGTGGC 1642

QY 584 GATTGGGGCGGACACGTCACTCCCTGGGTCAAAGAGA 620
Db 1643 AAAAGGCTCAAGGGCGCGCCCGCCACGCTTCGAAGGGA 1679

```

```

RESULT 14
PCT-US95-07180-1
; Sequence 1, Application PC/TUS9507180
; GENERAL INFORMATION:
; APPLICANT: LI, YI

```

```

; APPLICANT: GOCAYNE, JEANINE D
; APPLICANT: RUBEN, STEVEN M
; TITLE OF INVENTION: G-PROTEIN RECEPTOR HIBEB69
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
; ADDRESSEE: STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NJ
; COUNTRY: US
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07180
; FILING DATE: 06-JUNE-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 30,073
; REFERENCE/DOCKET NUMBER: 325800-366
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2453 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 548..1564
; PCT-US95-07180-1

```

```

Query Match 1.9%; Score 43.4; DB 5; Length 2453;
Best Local Similarity 47.3%; Pred. No. 0.054;
Matches 131; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

QY 344 CGTGTGGTCGACGTGGTCTCGACTGCGTCTATCCCTTACATCGACGACCCCAAGGACCGC 403
Db 1250 CGCATGATCGCCATAGTGTGGCCATCTTCTGGTCTGTTGTCCTTACACGTCAAC 1309

QY 404 GACGCGGTTTCCAGGTTGTGCGAGCTGTGTAAGAGCTCGACTCGCTACCCGCAAGCAC 463
Db 1310 CGCTCCGTCTAGTGTGCACTACCGGACGCCATGGGGCTCTTGGCCACCCAGCGCATC 1369

QY 464 GTCACATCGCGCTGTGTACACACCCCGGCTCGCTCGCGCGGCTTCCCGCAC 523
Db 1370 CTGGCCCTGGCAACCGCATACCTCTGCTTACACGCTTCAAGGGGCACTCGACCC 1429

QY 524 CTCGAGTCGCTCAAGCTCAAGGGCAAGCCCGAGCGCAATGTTCACCTTGATACCCGAG 583
Db 1430 ATCATGTATTCTTCTGTGCTGAGAAGTTCGCCACGCGCTGTGCAACTTGTCTGTGGC 1489

QY 584 GATTGGGGCGGACACGTCACTCCCTGGGTCAAAGAGA 620
Db 1490 AAAAGGCTCAAGGGCGCGCCCGCCACGCTTCGAAGGGA 1526

```

```

RESULT 15
US-08-947-823-1/c
; Sequence 1, Application US/08947823
; Patent No. 6114605
; GENERAL INFORMATION:
; APPLICANT: Williamson, Valerie M.
; APPLICANT: Kaloshian, Isghouhi
; APPLICANT: Yaghoobi, Jafar

```

Tue Apr 20 10:32:12 2004

APPLICANT: Bodeau, John  
APPLICANT: Milligan, Stephen  
TITLE OF INVENTION: Procedures and Materials for Conferring  
TITLE OF INVENTION: Pest Resistance in Plants  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/947,823  
FILING DATE: 09-OCT-1997  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US97/18802  
FILING DATE: 09-OCT-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/028,191  
FILING DATE: 10-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 023070-070210US  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 51952 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-947-823-1

Query Match 1.9%; Score 43.2; DB 3; Length 51952;  
Best Local Similarity 46.5%; Pred. No. 0.36; Indels 21; Gaps 3;  
Matches 264; Conservative 0; Mismatches 283;  
QY 1272 ATGTATTCGAGATAGAGGTTTAGAGTTCTTGGTGGTTGTTGTAAGAGGCTAAAGGC 1331  
DB 50449 ATTATATTGAAGATAGCGTCTTGAGGATTCGCCACACTTGTAAAGAACTTCAAGAGC 50390  
QY 1332 TTAGGATTCGAAAGGGCGGATGATCAAGGAATGGAGGATGAAGAGTACTGTGTCCC 1391  
DB 50389 TTAGGCT-----GTTTCCTTTTGATCTATTGTCTCCAGCCTAATGTATCTTTGACAG 50336  
QY 1392 ATAGAGGCTAATAGCCTTGTACAGGCGTGTTCAGAGCTTGAATACATGGCTGTTTATG 1451  
DB 50335 AGCAGGCGCTTGTAGTGTCTCAATGGGCTGCGCTAAGCTTCAGTCACTTTTACTTCT 50276  
QY 1452 TGTCTCATATTACAAATGCATCTCTGGAACATATTGGAACCTCAC-----TTGAAGAAC 1505  
DB 50275 GCGGCCAATGCAAAATAGCGCTTAGTTAGTATTGCNAGGACCGTTCTAACATGATCC 50216  
QY 1506 TCTGTGATTTTCGCTTGTCTGTTGACCATGAAGAGAGATACTGATTCGCACTTG 1565  
DB 50215 GATTTGTTTGTGTATTATCGAGCCTCAAACTCCTGACTACTTAATCTTTGAACCACTTG 50156  
QY 1566 ACAAATGGGGTGAAGGCTCTACTGAGGGCTGTGACAGCTGAGGAGATTTGCTCTATATC 1625  
DB 50155 ATACTGTTTTGGGGCCATTTGTGCAACACTGCAAGGAAGTGCAGCGACTTTCTCTTT--- 50099  
QY 1626 TCAGGCGTGGCGGTTTGCATCTGATGTAGGCGCTTGGTTACATTGGACAATACAGTCCAAATG 1685

Db 50098 -----CTGGTCTCCTTACAGATCGTGTGTTTGTAGTACATCGGGGTCCATGCTAAGAGT 50045  
QY 1686 TGAGATGGATGCTGCTTGGTTATGTGGGGAGTCTGATGCAAGGCTTTTGGAGTTGCTTA 1745  
Db 50044 TAGAGATGCTTTTCCTTAGCTTTTGGGGAGATAGTATATAGGCTTCTCTATGTTCTCT 49985  
QY 1746 AGGGGTGTCTTCTTCTTCAAACTTGAATGAGAGGGTGTATTATTTTTCAGTGAAAGCTG 1805  
Db 49984 CTGGTTTTCAGAGGCTTCCGTAAGTTGGAGACTAGAGACTGCCCTTTTGGCGATGAGGCTC 49925  
QY 1806 CACTTGTCTGGCTGCAACACACATTGAC 1833  
Db 49924 TGTGGCTAATGCTGCAAGCTGGAGAC 49897

Search completed: April 19, 2004, 22:59:27  
Job time : 166 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 20:29:51 ; Search time 1487 Seconds

(without alignments)  
6884.896 Million cell updates/sec

Title: US-10-009-791-21

Perfect score: 2288

Sequence: 1 gcacgagccacacgttaca.....acatttcatttgaat 2288

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2890132 seqs, 2237290429 residues

Total number of hits satisfying chosen parameters: 5780264

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA: \*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq: \*  
2: /cgn2\_6/ptodata/2/pubpna/FCI\_NEW\_PUB.seq: \*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq: \*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq: \*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq: \*  
6: /cgn2\_6/ptodata/2/pubpna/PCUUS\_PUBCOMB.seq: \*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq: \*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq: \*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq: \*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq: \*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq: \*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq: \*  
13: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq: \*  
14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq: \*  
15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq: \*  
16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq: \*  
17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq: \*  
18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq: \*  
19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2054	89.8	2305	13	US-10-424-599-141824
2	1999	87.4	2039	13	US-10-425-114-11061
3	1982	86.6	2011	13	US-10-425-114-10724
4	609.8	26.7	2178	16	US-10-310-154-152
5	561.2	24.5	2419	16	US-10-310-154-151
6	409.8	17.9	1503	13	US-10-425-114-32228
7	403.8	17.6	1325	13	US-10-425-114-8867
8	394.8	17.3	1469	13	US-10-425-114-33936
9	323	14.1	669	13	US-10-424-599-125710
10	311.6	13.6	744	13	US-10-424-599-125711
11	290.4	12.7	829	13	US-10-424-599-317
12	279.8	12.2	428	13	US-10-424-599-44612
13	227.2	9.9	719	9	US-09-770-149-161
14	220.2	9.6	449	13	US-10-424-599-21331

15	218.6	9.6	789	13	US-10-424-599-21332	Sequence 21332, A
16	199.2	8.7	477	13	US-10-424-599-27395	Sequence 27395, A
17	193.8	8.5	212	13	US-10-424-599-86569	Sequence 86569, A
18	185.8	8.1	350	13	US-10-424-599-38104	Sequence 38104, A
19	179.6	7.8	447	13	US-10-424-599-79984	Sequence 79984, A
20	166.4	7.3	417	13	US-10-424-599-39506	Sequence 39506, A
21	160.2	7.0	361	13	US-10-424-599-32267	Sequence 32267, A
22	125.4	5.5	1142	13	US-10-424-599-97745	Sequence 97745, A
23	124.4	5.4	283	9	US-09-294-0938-1569	Sequence 1569, Ap
24	112.2	4.9	521	13	US-10-424-599-28814	Sequence 28814, A
25	105.8	4.6	712	13	US-10-424-599-5625	Sequence 5625, Ap
26	105.6	4.6	711	13	US-10-424-599-34663	Sequence 34663, A
27	99.6	4.4	2793	13	US-10-424-599-102884	Sequence 102884, A
28	97.2	4.2	2465	13	US-10-425-114-26663	Sequence 26663, A
29	97	4.2	37	9	US-09-969-373-1044	Sequence 1044, Ap
30	96	4.2	2345	16	US-10-310-154-287	Sequence 287, App
31	86.2	3.8	682	13	US-10-424-599-93112	Sequence 93112, A
32	82.8	3.6	274	9	US-09-923-876-1583	Sequence 1583, Ap
33	82.8	3.6	274	11	US-09-923-876-1583	Sequence 1583, Ap
34	81.6	3.6	3433	13	US-10-424-599-91753	Sequence 91753, A
35	75.6	3.3	2693	13	US-10-424-599-69165	Sequence 69165, A
36	70	3.1	269	13	US-10-424-599-130155	Sequence 130155, A
37	68.6	3.0	1236	13	US-10-424-599-34657	Sequence 34657, A
38	67	2.9	1808	13	US-10-424-599-4935	Sequence 4935, Ap
39	65	2.8	302	13	US-10-424-599-83717	Sequence 83717, A
40	58.2	2.5	254	13	US-10-424-599-45283	Sequence 45283, A
41	52.6	2.3	2436	15	US-10-156-761-7456	Sequence 7456, Ap
42	52.6	2.3	9025608	15	US-10-156-761-1	Sequence 1, Appli
43	52.4	2.3	3147	13	US-10-424-599-4932	Sequence 4932, Ap
44	51	2.2	343	13	US-10-424-599-65547	Sequence 65547, A
45	50.6	2.2	327	13	US-10-424-599-137620	Sequence 137620, A

ALIGNMENTS

RESULT 1  
US-10-424-599-141824  
; Sequence 141824, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 141824  
; LENGTH: 2305  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_99079C.1  
US-10-424-599-141824

Query Match	89.8%	Score	2054;	DB	13;	Length	2305;
Best Local Similarity	99.5%	Pred. No.	0;				
Matches	2060;	Conservative	0;	Mismatches	10;	Indels	0;
Gaps	0;						
Qy	219	TTTCTCCTCATCATCCCAATACCATCCCTTTTGGCCTTGGCAACCAAACTCTG	278				
Db	191	TTCTTCTCCTCATCATCATCCCAATACCATCCCTTTTGGCCTTGGCAACCAAACTCTG	250				
Qy	279	CACCTTTCTTTTTCACCTCTCGATCCATATATGACGAGGAGCAACGTCGCGA	338				
Db	251	CACCTTTCTTTTTCACCTCTCGATCCATATATGACGAGGAGCAACGTCGCGA	310				
Qy	339	AGACACGTGTGGTCGAGTGGTCTCTCGACCTCGCTCATCCCTTACATCGACGCCCAAGG	398				



; CURRENT FILING DATE: 2003-04-28									
; NUMBER OF SEQ ID NOS: 73128									
; SEQ ID NO 11061									
; LENGTH: 2039									
; TYPE: DNA									
; ORGANISM: Glycine max									
; FEATURE:									
; OTHER INFORMATION: Clone ID: 701002452_FLI									
US-10-425-114-11061									
Query Match 87.4%; Score 1999; DB 13; Length 2039;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 2010; Conservative 0; Mismatches 0; Indels 1; Gaps 1;									
QY	278	GCACCTTTCTTTCTCACTCTCAGTCTCCGATCAATATGACGAGAACGGAACGTGGCGG	337						
DB	1	GCACCTTTCTTTTCACTCTCAGTCTCCGATCCAAATATGACGAGAACGGAACGTGGCGG	60						
QY	338	AAGACAGCTGTGTCGACGTGGTCTCGACTGGCTCATCTCCCTTACATCGACGCCCAAG	397						
DB	61	AAGACAGCTGTGTCGACGTGGTCTCGACTGGCTCATCTCCCTTACATCGACGCCCAAG	120						
QY	398	GACCGGACGCCGTTTCCAGGTGTGTCGACGTGGTGAAGAGCTCGACTCGCTCACCGGC	457						
DB	121	GACCGGACGCCGTTTCCAGGTGTGTCGACGTGGTGAAGAGCTCGACTCGCTCACCGGC	180						
QY	458	AAGACAGTCAACATCGGCTCTGCTACACACACCCCGGCTCGCTCCGCGCGCGCTTC	517						
DB	181	AAGACAGTCAACATCGGCTCTGCTACACACACCCCGGCTCGCTCCGCGCGCGCTTC	240						
QY	518	CCGCACTCGAGTGGTCAAGCTCAAGGCAAGCCCGAGCGCGCAATGTTCAACTTGATA	577						
DB	241	CCGCACTCGAGTGGTCAAGCTCAAGGCAAGCCCGAGCGCGCAATGTTCAACTTGATA	300						
QY	578	CCGAGAGATTGGGCGGACAGTCACTCCCTGGGTCAAGAGATTTCTCAGTACTCGAT	637						
DB	301	CCGAGAGATTGGGCGGACAGTCACTCCCTGGGTCAAGAGATTTCTCAGTACTCGAT	360						
QY	638	TGCCTCAAGAGCTCCACTCCGCGCATGATGTCAGAGATTCGATCTTCAAGATCTC	697						
DB	361	TGCCTCAAGAGCTCCACTCCGCGCATGATGTCAGAGATTCGATCTTCAAGATCTC	420						
QY	698	GCTCGTGAACCGGCTCAAGCTCTCAAGCTTGAAGTGTGCAAGTGTCCGGTTTCAAC	757						
DB	421	GCTCGTGAACCGGCTCAAGCTCTCAAGCTTGAAGTGTGCAAGTGTCCGGTTTCAAC	480						
QY	758	ACCGATGGTCTTTTCCATATCGTGGCTTTGCAAGATTTAAGAGTCTGTTTTTGAG	817						
DB	481	ACCGATGGTCTTTTCCATATCGTGGCTTTGCAAGATTTAAGAGTCTGTTTTTGAG	540						
QY	818	GAAGCTCAATCTTGAAGAGCGGAGATGGCTACACGAGTTCGTTTGAATAATACA	877						
DB	541	GAAGCTCAATCTTGAAGAGCGGAGATGGCTACACGAGTTCGTTTGAATAATACA	600						
QY	878	GTTCTTGAGACTCTCAATTTTACTTGAAGACATGCTGTGTTGAAGATTGAGACCTT	937						
DB	601	GTTCTTGAGACTCTCAATTTTACTTGAAGACATGCTGTGTTGAAGATTGAGACCTT	660						
QY	938	GAACTTTTGAAGTAAATTTGCCCACTTGTGCTGTGAACTTACTGCTGTGAATA	997						
DB	661	GAACTTTTGAAGTAAATTTGCCCACTTGTGCTGTGAACTTACTGCTGTGAATA	720						
QY	998	CTGGATCTGTGAACTTCTTTAAGCATGCTCTGGCTGGAGAGTGTGTTGGAGGCACC	1057						
DB	721	CTGGATCTGTGAACTTCTTTAAGCATGCTCTGGCTGGAGAGTGTGTTGGAGGCACC	780						
QY	1058	TACACGAGGAAACCAAGATACCTGCTATATATCATTTACAGCAAGATTATGTCATG	1117						
DB	781	TACACGAGGAAACCAAGATACCTGCTATATATCATTTACAGCAAGATTATGTCATG	840						
QY	1118	GGTTTAAACATATTTGGAAGAAATGAGTTGCCCATGTTGTTTATGTTGACCGCTACTA	1177						
DB	841	GGTTTAAACATATTTGGAAGAAATGAGTTGCCCATGTTGTTTATGTTGACCGCTACT-	899						

QY	1178	AAAAATTGGATCTCTCTATGCAATGCTAGACACGGAGGATCATTTGATGTTAAATCCAA	1237						
DB	900	AAAAATTGGATCTCTCTATGCAATGCTAGACACGGAGGATCATTTGATGTTAAATCCAA	959						
QY	1238	AGGTGTCGAAATCTGGAAGTCTTGGACAAAGAAATGTAATGGAGATAGAGGTTAGAG	1297						
DB	960	AGGTGTCGAAATCTGGAAGTCTTGGACAAAGAAATGTAATGGAGATAGAGGTTAGAG	1019						
QY	1298	GTTCTTGGTCTGTTTGAAGAGGCTTAAAGAGCTTAGGATTGAAAGGGCGATGATGAT	1357						
DB	1020	GTTCTTGGTCTGTTTGAAGAGGCTTAAAGAGCTTAGGATTGAAAGGGCGATGATGAT	1079						
QY	1358	CAAGGAATCGAAGATGAAGAGGTACTGTGTCCTCATAGAGGGCTAATAGCTTGTACAG	1417						
DB	1080	CAAGGAATCGAAGATGAAGAGGTACTGTGTCCTCATAGAGGGCTAATAGCTTGTACAG	1139						
QY	1418	GGCTGTTCAAGGCTTGAATACATGCTGTTTATGCTCTGATATTACAAATGCATCTCTG	1477						
DB	1140	GGCTGTTCAAGGCTTGAATACATGCTGTTTATGCTCTGATATTACAAATGCATCTCTG	1199						
QY	1478	GAACATATTGGAACTCACTTGAAGAACTCTGTGATTTTTCGCTTGTGTTGCTTGACCAT	1537						
DB	1200	GAACATATTGGAACTCACTTGAAGAACTCTGTGATTTTTCGCTTGTGTTGCTTGACCAT	1259						
QY	1538	GAAGAGAAATAACTGATTTGCCACTTGAACAATGGGGTGAAGGCTCTACTGAGGGGCTGT	1597						
DB	1260	GAAGAGAAATAACTGATTTGCCACTTGAACAATGGGGTGAAGGCTCTACTGAGGGGCTGT	1319						
QY	1598	GACAGCTGAGGAGATTTGCTCTATATCTCAGCGCTGGGGGTTGACTGATGAGGCTT	1657						
DB	1320	GACAGCTGAGGAGATTTGCTCTATATCTCAGCGCTGGGGGTTGACTGATGAGGCTT	1379						
QY	1658	GGTTACATTTGGACAAATACAGTCCAAATGTGAGATGGATGCTGTTGTTATGTTGGGAG	1717						
DB	1380	GGTTACATTTGGACAAATACAGTCCAAATGTGAGATGGATGCTGTTGTTATGTTGGGAG	1439						
QY	1718	TCTGATGACGGGCTTTTGGAGTTCGCTTAAGGGGTTGCTAGTCTTCAAGAACTTGAATG	1777						
DB	1440	TCTGATGACGGGCTTTTGGAGTTCGCTTAAGGGGTTGCTAGTCTTCAAGAACTTGAATG	1499						
QY	1778	AGAGGTGTTTATTTTTCAGTGAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1837						
DB	1500	AGAGGTGTTTATTTTTCAGTGAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1559						
QY	1838	CTTAGGTACTTGGGTGCAAGGTTTATGGTGTATCTCCATCTGACGATGCTTTTGGTA	1897						
DB	1560	CTTAGGTACTTGGGTGCAAGGTTTATGGTGTATCTCCATCTGACGATGCTTTTGGTA	1619						
QY	1898	ATGGCTCGACCCCTTTTGGAACTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1957						
DB	1620	ATGGCTCGACCCCTTTTGGAACTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1679						
QY	1958	AATCCAGATGAGACTGTAGTTTGGAGCATCTGCTCATATTTCTTGATATTTTCTCTT	2017						
DB	1680	AATCCAGATGAGACTGTAGTTTGGAGCATCTGCTCATATTTCTTGATATTTTCTCTT	1739						
QY	2018	GCAGGCGAGAGATCAGATTTTCCAGATACTGTGTGCTTGGACACCTGCGCATGCGTT	2077						
DB	1740	GCAGGCGAGAGATCAGATTTTCCAGATACTGTGTGCTTGGACACCTGCGCATGCGTT	1799						
QY	2078	GATACCTAGAGCCAGAGCTGTGTATATATACAGATTTCTTTTGTGTTTCTCTCTCCCT	2137						
DB	1800	GATACCTAGAGCCAGAGCTGTGTATATATACAGATTTCTTTTGTGTTTCTCTCTCCCT	1859						
QY	2138	TTTATATGCTGTTTCTATGTTTCTGCTCTATTTAGTCTTCAATTTAGACAATAGTCTTG	2197						
DB	1860	TTTATATGCTGTTTCTATGTTTCTGCTCTATTTAGTCTTCAATTTAGACAATAGTCTTG	1919						
QY	2198	TAATAAGCTGTTGTTTCTATTTGAATTTCTGAAAGCTTCCCTTAAACGCTATGCTGCC	2257						
DB	1920	TAATAAGCTGTTGTTTCTATTTGAATTTCTGAAAGCTTCCCTTAAACGCTATGCTGCC	1979						





1987 TCCTGCTCATATCTTGCATATATCTCTTGCAGGCGAGATCAGATTTCCAGATAC 2046  
1681 TCCTGCTCATATCTTGCATATATCTCTTGCAGGCGAGATCAGATTTCCAGATAC 1740  
2047 TGTGTGCTTTTGGACACTGCCACATCGTTGATACCTAGAGGCGAGCTGTGTATATA 2106  
1741 TGTGTGCTTTTGGACACTGCCACATCGTTGATACCTAGAGGCGAGCTGTGTATATA 1800  
2107 TACAGATTTCTTTTGTGTTTCTTCTCCCTTTCAATGCTGTTTCTATGTTCTGCTCT 2166  
1801 TACAGATTTCTTTTGTGTTTCTTCTCCCTTTCAATGCTGTTTCTATGTTCTGCTCT 1860  
2167 ATTGTGATGATTTAGACAAATAGTCTTGTATATAGCCCTGTTTCTATTTGAAATTC 2226  
1861 ATTGTGATGATTTAGACAAATAGTCTTGTATATAGCCCTGTTTCTATTTGAAATTC 1920  
2227 TGAACGCTTCCCTTAAAGCTATTTGGCTCCCTTAAACAACTGAACATTTCTCAATTTGTGA 2286  
1921 TGAACGCTTCCCTTAAAGCTATTTGGCTCCCTTAAACAACTGAACATTTCTCAATTTGTGA 1980  
2287 AT 2288  
1981 AT 1982

RESULT 4

US-10-310-154-152  
; Sequence 152, Application US/10310154  
; Publication No. US20030233670A1  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D  
; APPLICANT: Chomet, Paul S.  
; APPLICANT: Adams, Thomas H  
; APPLICANT: Ruff, Thomas G.  
; APPLICANT: Agarwal, Ameeta K.  
; APPLICANT: Ahrens, Jeffrey E.  
; APPLICANT: Ball, James A.  
; APPLICANT: Banu, G.  
; APPLICANT: Bell, Erin  
; APPLICANT: Boddupalli, Raghava  
; APPLICANT: Deikman, Jill  
; APPLICANT: Deng, Molian  
; APPLICANT: Duff, Stephen M.  
; APPLICANT: Galligan, Meghan M.  
; APPLICANT: Hinchey, Brenda S.  
; APPLICANT: Huang, Shihshieh  
; APPLICANT: Johnson, G. Richard  
; APPLICANT: Jung, Vincent  
; APPLICANT: Kretzmer, Keith A.  
; APPLICANT: Laccetti, Lucille B.  
; APPLICANT: Lai, Chao-Qiang  
; APPLICANT: Lee, Gary  
; APPLICANT: Lin, Jie-Yi  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Lu, Bin  
; APPLICANT: Luethy, Michael M.  
; APPLICANT: Lund, Adrian  
; APPLICANT: Madson, Linda L.  
; APPLICANT: Malloy, Kathleen A.  
; APPLICANT: McKiel, Christine L.  
; APPLICANT: Miller, Philip W.  
; APPLICANT: Padmavathi, Manchikanti  
; APPLICANT: Parnell, Laurence D.  
; APPLICANT: Start, William G.  
; APPLICANT: Tennesen, Dan  
; APPLICANT: Vidya, K.R.  
; APPLICANT: Wang, Haiyun  
; APPLICANT: Xin, Zhanguo  
; APPLICANT: Xu, Nanfei  
; APPLICANT: Yang, Chunzhi  
; APPLICANT: Zeng, Xiaoping  
; APPLICANT: Zhang, Qiang

APPLICANT: Zhao, Yajuan  
APPLICANT: Zhou, Li  
TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants  
FILE REFERENCE: 38-15(52796)B  
CURRENT APPLICATION NUMBER: US/10/310,154  
CURRENT FILING DATE: 2002-12-04  
PRIOR APPLICATION NUMBER: 60/337,358  
PRIOR FILING DATE: 2001-12-04  
NUMBER OF SEQ ID NOS: 736  
SEQ ID NO 152  
LENGTH: 2178  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (130)..(1923)  
OTHER INFORMATION:  
US-10-310-154-152  
Query Match 26.7%; Score 609.8; DB 16; Length 2178;  
Best Local Similarity 61.9%; Pred. No. 1.5e-176;  
Matches 1076; Conservative 0; Mismatches 642; Indels 21; Gaps 6;  
Qy 345 GTGTGCTGAGCTGCTCTGAGCTGCTCATCCCTTATCATCGACGCCCAAGGACCGCG 404  
Db 191 GCGTCCCGGAGGCGCTGCACCTCGTGTGGCTACATGAGACGCCCGGAGACCGGG 250  
Qy 405 ACGCGGTTTCCAGGCTGTGTCGACGCTGTGTACGAGCTCGACTCGCTCAACCGCAAGCACG 464  
Db 251 AGGCGGCTGCTGCTGTGCGGCTCTGGCACCGCATCGACGCTCTCGGCGAAGCACG 310  
Qy 465 TCACCATGCGGCTCTGTCTACACCAACCGCGGCTGCGCTCCGCGGCTTCCGCGACC 524  
Db 311 TCACCGTGGGCTTCTGTCTAGCGCGTGGAGCCCGCGGCTGCTCGCCAGGTTCCGCGCGC 370  
Qy 525 TCGAGTCGCTCAAGCTCAAGGGCAAGCCCGAGCGCAATGTTCAACTTGTATACCCGAGG 584  
Db 371 TCGAGTCGCTCGCGCTCAAGGGGAGGCCACCGCGCGCATATGACGGGCTCATCCCGAAG 430  
Qy 585 ATTGGGCGGACACGCTCACTCCCTGGGTCAAAGAGATTTCTCAGTACTTCGATTCGCTCA 644  
Db 431 ACTTTGGGCGCTAGCGGCGGCTGCTGGTTCGCGAGCTCGCGCGGCCCTTGACTGCTCA 490  
Qy 645 AGAGCTCCACTTCGCGCGCATGATGTCAAGGATTCGATCTTCAGATCTTCGCTCGTG 704  
Db 491 AGGCGCTCCACTCGCTGCTGATGACCGTCAACCGAGGACATCGCGCTGCTGCTCCAG 550  
Qy 705 ACCGCGGTGAGTCTTCAAGCTCTCAAGCTTGAAGTGTCCGCTTTCACACCGGATG 764  
Db 551 CGCGCGGCCACATGCTACAGGCGCTCAAGCTCGACAGTGTCCGCTTCTCAACAGAGC 610  
Qy 765 GTCTTTTCCATATCGTCTGCTTTTTCAGAGATTTAAGAGTCTTGTGTTTGGAGGAAAGCT 824  
Db 611 CCCTCGGCTCGTCCGCGGCTCTCTGAGAACCTTGTTCCTGGAGAAATGTA 670  
Qy 825 CAATTTTGAAGAGACGAGAAATGCTGCTACAGACTTGTCTTTGAATATACAGTTCTTG 884  
Db 671 TAAATTGAGGATGAAGGAGATGAATGGCTCCATGAATCTGCTGCTCAACAATCTGTTCTGG 730  
Qy 885 AGACTCTCAATTTTACTTGTAGACAGATGCTGTTGTGAAGATTTGAGGACCTTGAACCTTT 944  
Db 731 TGACACTGAACCTTTTACATGACAGAACT---CAAGTGGAGGCTGCTGATCTGGAGCTTC 787  
Qy 945 TAGCTAAAAATTCGCCCAACTTAGTGTCTGTGAAACTTTACTGACTGTGAAATCTGGATC 1004  
Db 788 TTGCAAGAGACTGTAAATCATTTGATTTCTTTGAAGATGGTGTGACTGTGATCTTTTTCAGATC 847  
Qy 1005 TTGTGAACCTTTTAAAGCATGCTCTGCGCTGGAAGAGTTTGTGGAGGACCTTCAACG 1064  
Db 848 TGATTGGGTTTTTCCAAACCTCCAAAGCATTTGCAAGAAATTCGCGGTGTGCTTTTGTG 907  
Qy 1065 AGGAACACAGA-----AGATACTCTGCTATATCATTTACAGCAAGTATATGTCGAT-- 1116

Db 908 AAGTTGAGAGTACACCAAGTACGAAAGGTCAFTTTCCACCTAGGCTATGCTTCTTAG 967  
Qy 1117 -GGGTTTAAATATTTGGAAGATGAGTTGGCCATTGTTTCATCTTTGACCGGTAC 1175  
Db 968 GGGGCTTACCTTCATGGTAAACAGAGTGGCCGTTATCTTTCCGATTTCCAAATGC 1027  
Qy 1176 TAAAAAATGGATCTCCTCTATGCAATGCTAGACAGGAGATCATTTGATGTTATATCC 1235  
Db 1028 TTAAGAACTGAGCTTGAGTTCACTTTCTCCACCTGAAGACCACTGTGAGCTTATG 1087  
Qy 1236 AAAGGTGTCAAATCTGGAAGTCTCTGAGACAGGAATGTAATTTGGAGATAGAGGTTAG 1295  
Db 1088 CTAATATGCCCACTAAGTCTTCGAGGTGAGGAATGTGATTTGGGATAGAGACTAG 1147  
Qy 1296 AGTTCTTGCTGTTGTTGTAAGAGGCTTAAAGGCTTAGGATTCAGAGGGCGGATGATG 1355  
Db 1148 AAGTTGTTGCTGTACATGCAAGACTACGAGGCTCAGAGGCTCAGAAATGAGCGTGCAGCATG 1207  
Qy 1356 ATCAAGGAATGGAGATGAAGAGTACTGTGTCCCATAGAGGGCTAATAGCCTTGTAC 1415  
Db 1208 ATCTGGCCCAAGAGAGCAGGAGG--AGTCTCTCAGATAGTTGACAGCCGTAGCCG 1264  
Qy 1416 AGGCTGTTCAGAGTTGAAATACATGCTGTTTATGCTGTGATATTAATAAATGATCTC 1475  
Db 1265 TTGGCTGCCGGAATGAGTACATAGCTGCTATGATCTGATATCAAAATGGGGCGC 1324  
Qy 1536 ATGAAGAAGATAACTGATTTGCCACTGTGAATGGGGTGAGGGCTCTACTGAGGGCT 1595  
Db 1385 AACAGAAGATAACAGACCTGCCATTTGACAGGGGCTCGGCTCTACTGAGGAAT 1444  
Qy 1596 GTGCAAGCTGAGGATTTGCTCTATATCTCAGGGTGCGGGTTGACTGATGTAGGCC 1655  
Db 1445 GCGTCAAACTCGGAGGTTGCTCTCTACTGAGACCGGAGGGCTTTGAGAGTTAGGCC 1504  
Qy 1656 TTGGTTACATTTGGAACAATGAGTCCAAATGTGAGATGAGTGTCTGTTGTTATGTTGGG 1715  
Db 1505 TCGGTTACATTTGGAATATACAGTGGGAACATCCAATACATGCTGCTCGGCAAGTTGGTG 1564  
Qy 1716 AGTCTGATGACGGCTTTTGGAGTTCGCTTAAGGGGTGCTAGTCTTCAGAACTTGAA 1775  
Db 1565 AATCTGACAAAGGATGATCCAGTTTGCACCGGATGCAACCTCGGAGCTCGAGC 1624  
Qy 1776 TGAGAGGTGTTTATTTTTCAGTGACGTGCACTTCTGTGGCTGCAACACAAATGACTT 1835  
Db 1625 TGAGGGGCTG--CTGCTTACGAGGAGCTCTGGCAGTGGCCGTCTCCAGATGCCCT 1681  
Qy 1836 CTCTTAGGTACTTGTGGGTGCAAGTTATGGTGTATCTCAATCTGACGTGATCTTTTG 1895  
Db 1682 CGCTGAGGTACATATGSGTGCAGGGTACAGAGCTCTCGAAGGGCCAGGACCTCATGC 1741  
Qy 1896 TAATGGCTGACCTTTTGGAAATTTGAGTT---GATTCCTTCTAGAAAGTGTGCTACGA 1952  
Db 1742 TCATGGCCAGCGGTACTGGAACATTTGAATTCGGGCTCCCATTCGAGAGGCGCTTATC 1801  
Qy 1953 ATACCAATCCAGATGAGACTGTAGTTGTTGAGCATCTGCTCATATTTCTTGCATATTT 2012  
Db 1802 GGGTGTGAGGAGATGAGACGCTTGTGTTGACACCCATGCCAGGCTCTTGGTATTTACT 1861  
Qy 2013 CTCTTGAGGGCGAGATCAGATTTTCCAGATATGTTGCTGCTTTGGACACTGCCACA 2071  
Db 1862 CCCTTGTGGAAGAGGCGGAGTGCCTCTAGTGGTGGTGAATTTGACCTTCCGCTCA 1920

RESULT 5

US-10-310-154-151  
; Sequence 151, Application US/10310154  
; Publication No. US20030233670A1  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D

; APPLICANT: Chomet, Paul S.  
; APPLICANT: Adams, Thomas H  
; APPLICANT: Ruff, Thomas G.  
; APPLICANT: Agarwal, Ameeta K.  
; APPLICANT: Ahrens, Jeffrey E.  
; APPLICANT: Ball, James A.  
; APPLICANT: Banu, G.  
; APPLICANT: Bell, Erin  
; APPLICANT: Boddupalli, Raghava  
; APPLICANT: Deikman, Jill  
; APPLICANT: Deng, Molian  
; APPLICANT: Dong, Jinzhao  
; APPLICANT: Duff, Stephen M.  
; APPLICANT: Galligan, Meghan M.  
; APPLICANT: Hinchey, Brenda S.  
; APPLICANT: Huang, Shihshieh  
; APPLICANT: Johnson, G. Richard  
; APPLICANT: Jung, Vincent  
; APPLICANT: Kretzmer, Keith A.  
; APPLICANT: Laccetti, Lucille B.  
; APPLICANT: Lai, Chao-Qiang  
; APPLICANT: Lee, Gary  
; APPLICANT: Lin, Jie-Yi  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Lu, Bin  
; APPLICANT: Luethy, Michael M.  
; APPLICANT: Lund, Adrian  
; APPLICANT: Madson, Linda L.  
; APPLICANT: Malloy, Kathleen A.  
; APPLICANT: McKiel, Christine L.  
; APPLICANT: Miller, Philip W.  
; APPLICANT: Padmavathi, Manohikanti  
; APPLICANT: Parnell, Laurence D.  
; APPLICANT: Start, William G.  
; APPLICANT: Tennesen, Dan  
; APPLICANT: Vidya, K.R.  
; APPLICANT: Wang, Haiyun  
; APPLICANT: Xin, Zhanquo  
; APPLICANT: Xu, Nanfei  
; APPLICANT: Yang, Chunzhi  
; APPLICANT: Zeng, Xiaoping  
; APPLICANT: Zhang, Qiang  
; APPLICANT: Zhao, Yajuan  
; APPLICANT: Zhou, Li  
; TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants  
; FILE REFERENCE: 38-15(52796)B  
; CURRENT APPLICATION NUMBER: US/10/310,154  
; CURRENT FILING DATE: 2002-12-04  
; PRIOR APPLICATION NUMBER: 60/337,358  
; PRIOR FILING DATE: 2001-12-04  
; NUMBER OF SEQ ID NOS: 736  
; SEQ ID NO 151  
; LENGTH: 2419  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (139)..(1911)  
; OTHER INFORMATION:  
US-10-310-154-151

Query Match 24.5%; Score 561.2; DB 16; Length 2419;  
Best Local Similarity 61.3%; Pred. No. 1.7e-161;  
Matches 1000; Conservative 0; Mismatches 603; Indels 27; Gaps 5;  
Qy 327 GGAACTGCGGAAGACACAGTGTGTCGACGTGCTCTCGACTGCGTTCATCCCTTATATCG 386  
Db 176 GGCGGGTGTGAGCTTTTGGGATCCCGGACACCGCGCTAGTGTATGGGGTATCGTG 235  
Qy 387 ACCACCCCAAGGACCGCGCGTTCCTCCAGTGTGTCCAGCTGTGTCCAGCTGTGTACGACT 446  
Db 236 AGGATCCCTGGGACCGCGGACCGCATCTCGTGTGTGTGCCGACACTGTTGCGCGGTGACG 295

QY	447	CGCTCACCCGCAAGACACGCTCACCATCTCGCGCTCTGCTACACCAACACCCCGGCTCGCCTCC	506
Db	296	CGCTCAGCCGCAAGACACGCTCACCGTCGCCATGCGCTACTCCACAAACCCGAGCGCCTGT	355
QY	507	GCGCGCGTTCCGCAACCTCGAGTCGCTCAAGCTCAAGGCAAGCCCGAGCGCCAACTGT	566
Db	356	TCGCGGCGCTCCCGTGCCTCGAGTCACCTCAAGCTCAAAAGCAAGCCCGCGCTCAATGT	415
QY	567	TCACCTTGATACCGAGGATTTGGGCGGACAGTCACTCCCTGGGCTCAAAAGAGATTTCTC	626
Db	416	TCACCTCACTCCGACGACTGGGCGGGTCTGCGTCGCCCTTGGATCCGACAGCTCTCGG	475
QY	627	AGTACTTCGATTTGCCTCAAGAGCCTCCACTTCGCGCGCATGATGTCAAGGATTCGGATC	686
Db	476	CCACCTTCACCTTCCCTCAAGAAGCTCCACTCGCGCAGGATGATAGTATCTCAACGGAAGATA	535
QY	687	TTCCAGAAATCTCGCTCGTGACCGCGGTCACGTGCTTCACGCTCTCAAGCTTGCACAGTCTC	746
Db	536	TCACCATCTCTTGTGCGTGCCAGCGCAATGCTGCTGCGCTGAAGCTCGACCGCTGCT	595
QY	747	CCGTTTTCACCAACCGATGTCCTTTCCATATCGGTGCTCTTTTGCAGAGTTTTAAAGTCT	806
Db	596	CCGCTTCTCCACTCCCTCCATCGCACTCGTCGCCCGTCTCTGCAGGAACCTGGAACAC	655
QY	807	TGTTTTTGGAGGAAAGCTCAATTTCTTGAAGGAC--GGAGAAATGGCTTACACGAGCTTG	863
Db	656	TTTTTCTGGAGAAAGACAAATTTGATGAGAAAGAAAATGATGATGGATCGCTGAGCTTG	715
QY	864	CTTTGAATATATACAGTTCTTGAGACTCTCAATTTTACTTGCACAGACATTCGTGTTGTA	923
Db	716	CTACGAGCAATTCCTGTTCTTGAGACACTGAATTTCTTTCTTAACAGA--TCTCAGGGCAT	772
QY	924	AGATTGAGGACCTTGAACCTTTTACTTAAATTTGCCCAACTTAGTGTCTGTGAAACTTGA	983
Db	773	CCCAGAGTATCTTACCTCTTGTGGCGCACTGTCAAAGATTGAAAACCTCTGAGATTA	832
QY	984	CTGACTGTGAAATACTGGATCTTGTGAACTCTTTTAAAGCATGCTCTGCGCTGGAAGAT	1043
Db	833	GTGAATGTTTCATGCCGATCTGGTCAGTTGTTCGAACTGCACAAACACTACAAGAT	892
QY	1044	TTTGTGGAGGCACCTACAACGAGAACCCAGAAATACTCTGCTATA-----	1090
Db	893	TCGCTGSGTGTCTCTTTGAAGAGCAGGGTCAACCTGTGGCAATGAGAAATTATGAAACT	952
QY	1091	--TCATTACCAAGAAAGTTATGTTCGATTGGTTTTAAACATATATTTGGAAGATAGTTGC	1148
Db	953	ACTATTTTCTCTCTCACTGCACCACTTGAGTTTGCTCTACATGSGAACAAATGATATGC	1012
QY	1149	CCATTGTCTCATGTTTSCAGCCGCTATAAAAAATTGGATCTCTCTATGCAATGCTAG	1208
Db	1013	AAATACTGTTTCCATATGCTACTGCATCTAAGAATTAGACCTTCAGTTTACATTCCTTT	1072
QY	1209	ACACGGAGGATCATTTGATGTTAATCCAAAGTCTCCAAATCTGGAAGTCTTTGAGACAA	1268
Db	1073	CCACAGAGGATCATTTGCAGATAGTTCAACGCTGCTCCAAATCTTGAAAAACCTTAGAGTG	1132
QY	1269	GGAAATGTAATTGGAGATAGGGTTTAGAGGTTCTTTGTCGTTGTTTAAAGAGGCTAAAA	1328
Db	1133	GGGATGTCTAGGGGATCGTGGACTACAAGTTTGTGCACAGCTGCACAGAAATTTGCATA	1192
QY	1329	GGCTTAGGATTTGAAAGGGCGCATGATGATCAAGG---ATGGAGGATGAAAGAGTACTG	1385
Db	1193	GGCTCAGAGTAGAGAGAGGAGATGATGATCAAGGAGTCTTTGAGGATGAAACAAGGTAGGA	1252
QY	1386	TGTCCCATAGAGGCTAATAGCCTTGTACAGGGCTGTTACAGGCTTGAATACATAGGCTG	1445
Db	1253	TTTCAAGGTGGGGTTGATGGCTATAGGCCAAGGCTGCCCTGAGTTGACATACCTGGGCGA	1312
QY	1446	TTTATGTGCTGATATTAACAAATGCACTCTCTGGAACATATTTGGAACCTCACTTTGAAGAAC	1505
Db	1313	TACATGATACAGACATTACAAATGACGCTTTTAGAGGCAGTTGTTACATGACGACAAAATC	1372
QY	1506	TCTGTGATTTTCGCTCTGTGTTGCTGACCATGAAAGAGAAAGATTAAGTATTTGCCACTTG	1565

Db	1373	TTAATGACTTCGGCCTTGCTCTTGATAGAGAGCACATATAACCGAATTGCCACTGG	1432
QY	1566	ACAATGGGTGAGGGCTCTACTGAGGGGCTGTGACAAGCTGAGGAGATTGGCTCTATATC	1625
Db	1433	ACAATGGGGTTCGTGCTTTAGAGGTTTGACACAACTACGGAGGTTTTGCATTTTATG	1492
QY	1626	TCAGGCGTGGCGGTGTGACTGATGATAGGCTTGGTTACATTGGACAATACAGTCCAAATG	1685
Db	1493	TGAGACTCTGGGGCCCTATCTGATGTTGGTCTTGGCCTATGTTGGAGAAATTTAGTAAGAGTA	1552
QY	1686	TGAGATGSGATGCTGCTTGTTGTTATGTGGGGAGTCTGATGACAGGCTTTTGGAGTTCGCTA	1745
Db	1553	TTCGTTTATATGTTGCTTGGTAAATGTTGGTGAATCTGATAATGGAATCATACAATTATCAA	1612
QY	1746	AGGGGTGTCCTAGTCTTCAGAACTTCGAAATGAGAGGGTCTTTTATTTTTTCAGTGAACGTG	1805
Db	1613	AAGGCTGCCAAGCTTGC AAAA ACTGAGGTGAGGGTGTGCTCTTT--AGTGAGCATG	1669
QY	1806	CAC TTGCTGTGGCTGCCAACACAAATTGACTTCTCTTAGGTACTGTGTGGTGCAAGGTTATG	1855
Db	1670	CTTTAGCTTTGGCTGCACTACAGCTTAAGTCACCTGAGGTATCTGTGGGTACAGGATTC	1729
QY	1866	GTGTATCTCAATCTGACAGTGAATCTTTTGGTAATGGCTCGACCCCTTTTGGAAACATTGAGT	1925
Db	1730	GGTCATCTCCACTGGAACTGATATTATGCAATGGTACGCCCCCTTCTGGAAACATTGAGT	1789
QY	1926	TGATTCCTTC	1935
Db	1790	ATAATGTTC	1799

RESULT 6  
 US-10-425-114-32228  
 ; Sequence 32228, Application US/10425114  
 ; Publication No. US20040034888A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jingdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Screen, Steven E  
 ; APPLICANT: Tabaska, Jack E  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53313)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,114  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO 32228  
 ; LENGTH: 1503  
 ; TYPE: DNA  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: UC-ZMFLB73276H05\_FLI  
 US-10-425-114-32228

	Query Match	17.9%;	Score 409.8;	DB 13;	Length 1503;
	Best Local Similarity	60.8%;	Pred. No. 7e-115;		
	Matches 780;	Conservative	0;	Mismatches 482;	Indels 21; Gaps 6
QY	801	GAGCTCTGTTTTTCGAGGAAAGCTCAATTC	TTGAGAAGGACGGAGAATGGCTTACACGAGC	860	
DB					
	1	GAACCTTGTCTCTGGAAGATGTATAT	TGAGGTGAGGAGTGAATGGCTCCATGAAC	60	
QY	861	TTCGCTTGAATAATAACGTTCTCTTGAGACTCTCAATTTT	TACTTGAACAGACATTCGCTGTTG	920	
DB	61	TGCGTGTCAACAAATCTGTCTCTGGTGACACTGAACTTTT	TACATGACAGAACT---	CAAAG 117	
QY	921	TGAAGATTGAGACCTTGAACCTTTT	TAGCTAAAAATGGCCCAACTTAGTGTCTGTGAAC	980	
DB	118	TGAGCCCTGCTGATCTGGAGCTCTTGTG	CAAGAACTGTAATCATTTGATTTCTTTGAAGA	177	

Tue Apr 20 10:32:12 2004

us-10-009-791-21.rnpb

981 TTACTGACTGTGAATACCTGATCTTGTGAACCTTCTTTAAGCATGCTCTGCGCTGGAAG 1040  
 Db TGGGTGACTGTGATCTTTCAGACTGATGGGTTTTCCTCAACCTCCAAAGCATTTGCAAG 237  
 1041 AGTTTTGTGAGGACCTACACAGAGAACAGAG-----AGATATCTTGTCTATATCAT 1094  
 Db AATTCGCGGTGGTGGCTTTTGAAGTGTGAGAGTACACCAAGTACGAAAGGTCAATTT 297  
 1095 TACAGCAAGTTATGTCGATT---GGGTTTAAATATATTTGGAAGAAATGAGTTGCCCA 1151  
 Db TCCACACCTAGCTATGCTCTTACGGGGTCTTACCTTCATGGGTGTAAGAGAGATGCCAG 357  
 1152 TTGTGTTCACTTTGCGACCGCTACTAAAAAATTTGGATCTCTCTATGCAATGCTAGACA 1211  
 Db TTATCTTTCGATTTCCACATGCTTAAGAAACTGGACTTCAGTTCACTTTCTCTCCACA 417  
 1212 CGAGGATCATGTTATGTTATCAAGGCTGTCCTCAAGTCTGGAAGTCTCTGAGACAGGA 1271  
 Db CTGAAGACCACTGTCAGCTTATGCTAAATGCCCAACCTAAGTGTCTCGAGGTGAGGA 477  
 1272 ATGTAATTTGGAGATAGAGGCTTAGAGGTTCTTTGGTCTGTTGTTAAGAGGCTAAAGGC 1331  
 Db ATGTAATTTGGAGATAGAGGCTTAGAGGTTCTTTGGTCTGTTGTTAAGAGGCTAAAGGC 537  
 1332 TTAGGATTTGAAGGGCGCATGATCAAGAAATGGAGATGGAAGGCTAGTGTCTCTCC 1391  
 Db TCAGAATTTGAGCGTGGCAGCATGATCTCTGGCCAAAGAGAGCAGGAGG---AGTCTCTC 594  
 1392 ATAGAGGCTATAGCTTGTTCACAGGCTGTTTCAGAGCTTCAATACATGCTGTTTATG 1451  
 Db AGATAGGTTGACAGCGGTAGCGGCTGCTGGCGAAGTACATAGTACATAGTGTCTCTATG 654  
 1452 TCTCTGATATTAACAATGATCTCTGGAACATTTGGAATCTCACTTGAAGAACCTCTGTG 1511  
 Db TATCTGATATCAAAATGGGCGCTGGATCCATTTGGCACTTCTGCAAGATCTCTATG 714  
 1512 ATTTGCGCTTGTGTTGTTGACATGAGAGAGATGATCTGTTGCGACCTTCTGCAAGATCT 1571  
 Db ATTTGCGCTTGTGTTGTTGACATGAGAGAGATGATCTGTTGCGACCTTCTGCAAGATCT 774  
 1572 GGGTGAAGGCTTACTGAGGCGGTGTGCAAGCTGAGAGATGTTGCTCTATATCTCAGCG 1631  
 Db GCGTCCGTGCTTACTGAGGAACTGGCTCAACTTGGAGGTTGCTTCTTCTACCTGAGAG 834  
 1632 GTGCGCGGTTGACTGATGTAGGCTTTGTTTACATGCAATACATGCTCAATGTGAGAT 1691  
 Db CGGAGGCGCTTTCAGACGTAGGCTCGGTTTACATTTGACATATACAGTGGGAAATCCCAAT 894  
 1692 GGATGCTGCTGTTTATGCGGGAGTCTGATGCGAGGCTTTTGGAGTTGCGTAAGGGGT 1751  
 Db ACATGCTGCTCGGCAAGCTTGTGTAATCTGACAACTGATTCAGTTTGGCAACGGGAT 954  
 1752 GTCTAGTCTTCAAGAACTTGAATAGAGGCTGTTTATTTTTCAGTGAACGTGCACTTG 1811  
 Db GCACCAACCTCGGAACTGAGCTGAGGCGCTG---CTGCTTTCAGGAGCGAGCTCTGG 1011  
 1812 CTGTGGCTGCAACACAAATGACATTTCTTATGTTTCTGTTGCGGTCGAAAGTTATGTTAT 1871  
 Db CAGTGGCGGTGCTCCAGATGCTCTGCTGAGTACATATGGGTGCGAGGCTTACAGAGCCT 1071  
 1872 CTCATCTGAGAGCTGATCTTTGTAATGCTCGACCTTTTGGAAACATTTGAGTT---GA 1928  
 Db CTCGAACCGGCGCAGGACCTTATGCTGATGCTGAGTACATATGGGTGCGAGGCTTACAGAGCCT 1131  
 1929 TTCTTCTTGAAGAGGTGCTACGAATACCAATCCAGATGAGACTGTAGTTGTTTTCAGCATC 1988  
 Db TCCCATCTCCGAGAGCGCTTATCGGTTGATGCGAGATGGACAGCTTGTGTTGACACCC 1191  
 1989 CTGCTCATATTTCTTGCATATTTCTCTTTGAGGCGAGAGATCATGATTTTCCAGATCTG 2048  
 Db ATGCCAGGCTCTTGGCTTATCTCTCTTGTGGAAGGAGGCGGAGCTGCGCTCTCAGTGGC 1251  
 2049 TTGTGCTTTGGACACTGCCACA 2071

1252 TGGTGACTTTGACCCCTGGCTCA 1274  
 Db  
 RESULT 7  
 US-10-425-114-8867  
 ; Sequence 8867, Application US/10425114  
 ; Publication No. US20040034888A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jingdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Screen, Steven E.  
 ; APPLICANT: Tabaska, Jack E.  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53313)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,114  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO 8867  
 ; LENGTH: 1325  
 ; TYPE: DNA  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: 700800510\_FLI  
 US-10-425-114-8867

Query Match 17.6%; Score 403.8; DB 13; Length 1325;  
 Best Local Similarity 60.2%; Pred. No. 4.6e-113; Indels 12; Gaps 3;  
 Matches 726; Conservative 0; Mismatches 467;  
 QY 923 AAGATTGAGGACCTTGAACCTTTTAGCTAAAAATTCGCCCAACTTAGTGTGTGTAACCTT 982  
 Db 32 AGGACAGTGTGCTTAGAGCTTCTTGCAGAGAGCTGCAAGTCACTGATTTCTCTTGAAGATT 91  
 QY 983 ACTGACTGTGAATACTGATCTTGTGAACTTCTTTAAGCATGCTCTGCGCTGGAAGAG 1042  
 Db 92 GCGGATTGTGATCTTTCAGATTTGATAGGGTTCTTCCAACTCTGCCACATCATCTGGAAGAA 151  
 QY 1043 TTTTGTGAGGACCTTACACAGGAGCAAGAA-----AGATATCTTCTCTATATCATTTA 1096  
 Db 152 TTTGTGAGGAGCAATTTAATGGCAGAGGGAACCTCACCAAGTATGGGATGTAAATTT 211  
 QY 1097 CCAGCAAACTTATGCTGATTTGGGTTTAAATATATTTGGAAGAAATGAGTTGCCATTGTG 1156  
 Db 212 CCATCAAGAAATATGCTCTTGGGACTTACTTTTCATGGGTGCAAAATGAAATGCTATTATA 271  
 QY 1157 TTTGATGTTTGCAGCGGTACTTAAATAAATTTGGATCTCTCTATGCAATGCTAGACAGGAG 1216  
 Db 272 TTTCTTTTCTGCGATCTTAAAGAGCTGGATTTTGCAGTACACTTTCTCTCACCGGAA 331  
 QY 1217 GATCATTTGATGTTATCCAAAGGCTGTCAAATCTGGAAGTCTCTTGAGACAGGAATGTA 1276  
 Db 332 GACCATTTGCCAGCTCAATGCAAAATGCCCAACTTACTAGTTCTCGGGTGAAGATGTG 391  
 QY 1277 ATTGGAGATAGAGGTTTAGAGGTTCTTGGTCTGTTGTTGTAAGAGGCTTAAAGGCTTAGG 1336  
 Db 392 ATTGGGATAGAGATTTAGAGTTTGTTCAGATACGTCGCAAGAGCTCCAAAGGCTCAGA 451  
 QY 1337 ATTGAAGGGGCGCATGATGATCAAGNAATGGAGGATGMAAGGTAAGTCTGTGCTCCATAGA 1396  
 Db 452 ATTGAGCGAGAGCAGATGAAGGAGGTGTCAGAGGATGMAAGGTAAGTCTGTGCTCCATAGA 511  
 QY 1397 GGGCTAATACCTTGTTCACAGGCTGTTTCAGAGCTTGAATATACATGCTGTTTATGTGCT 1456  
 Db 512 GGGTTCAGCTATAGCGGTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 571  
 QY 1457 GATATTAACAATGATCTCTGGAACATATTTGGAATCACTTGAAGAACCTTCTGTGATTT 1516  
 Db 572 GATATTAACAATGAGGCTTTGGAATCTATCGGACATTTCTGCAAAAAAATCTATGATCTC 631

QY	1517	CGCCTTGTGTTGCTTGACCATGAAGAGAAGATAACTGATTTGCCACTTGACAATGGGGTG	1576
Db	632		691
QY	1577	AGGGCTCTACTGAGGGCTGTGCAAGCTGAGAGATTTCCTCTATCTATCTCAGGCGTGCG	1636
Db	692		751
QY	1637	GGGTTGACCTGATCTAGCGCCTTGGTTACATTGGACAATACAGTCCAAATGTGAGATGGATG	1696
Db	752	GGGCTCTCAGATCGAGCTCTCGGCTACATTGGACAGTGCAGTGGAAACATTTCAATACATG	811
QY	1697	CTCGTTGGTTATGTGGGGAGTCTGATGCAAGGCTTTTGGAGTTCCGTAAGGGGTCTCCT	1756
Db	812	CTTCTCGGTAACTGGGGAAACCGATGATGGATTGATCAGTTTCGCTCTGGAGATGTGA	871
QY	1757	AGTCTTTCAGAAACTTGAATAGAGAGGTGTTTATTTTTCAGTGAAACGTGCACTTGCTGTG	1816
Db	872	AACCTGGAAAGCTTGAGCTCAGGAGCTGT--TGCTTCAGCGAGCGAGCTCTGGGCCCTC	928
QY	1817	GCATGCAACAATTTGACTTCTTTAGGTACTGTGGGTGCAAGGTATGGGTATCTTCCA	1876
Db	929	GCAATACTAAGTATGCCCTTCCCTGAGGTACGTATGGGTTCAGGGCTACAAGCGTCTCAA	988
QY	1877	TCTGGAGCTGATCTTTTGGTAAATGCGCTCGACCCCTTTTGGAAACATGTAGTTTGATTCCTTCT	1936
Db	989	ACAGGCGAGACCTTGATGCTCATGGCAGGCCCTTCTGGAAACATAGAGTTTACACCTCCC	1048
QY	1937	AG----AAAGGTGGCTACGAATACCAATCCAGATGAGACTGTAGTTGTTGAGCATCCTGCT	1993
Db	1049	AGTTCCCAAGATGCTGGTCCGGTTGATAGAGATGGGAAACCTTTCGCTAGATAGTCATGCT	1108
QY	1994	CATATTTCTTGATATATTCTCTTTCAGGGGAGAGATCAGATTTTTCAGATACCTTTGTG	2053
Db	1109	CAGATACCTGGCATACGGCTCTCTGCTGGTAAAGAGTTGGACTCGCCCACAATCCGCTGGTC	1168
QY	2054	CCTTTGGACACTGCCACATGCGTTGATACCTAGAGGCCAGAGCTGTGTATATATACCAGT	2113
Db	1169	ACTTTGATCTCTGCTGGTGAATAGACTGAGGGCCCGCTTTTGGTGCCCATGGAAACCGT	1228
QY	2114	TTTCT 2118	
Db	1229	TTCT 1233	

```

RESULT 8
US-10-425-114-33936
; Sequence 33936, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 33936
; LENGTH: 1469
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMELM017190F07_FLI
US-10-425-114-33936

```

Query Match 17.3%; Score 394.8; DB 13; Length 1469;  
Best Local Similarity 60.9%; Pred. No. 3e-110;  
Matches 753; Conservative 0; Mismatches 462; Indels 21; Gaps 6;

848	2y	TGGCTACAGAGCTTGGCTTTGAATAATACAGTTCCTTCGAGACTCTCAATTTTAACTTGACA	907
2	Db		
908	2y	GACATTCCTCTGTGTGAAGATTGAGGACCTTGAACCTTTTAGCTTAAAAATGCCCCAACTTA	967
62	Db		
968	2y	GTGTCGTGAACTTACTGACTGTGAAATACCTGATCTTGTGAACCTCTTTAAAGCATGCC	1027
119	Db		
1028	2y	TCCTCGCTGGAAGAGTTTGTGGAGGCACCTTACAACGAGGAAACAGAA	1081
179	Db		
1082	2y	TCTGCTATATCATTTACAGACAAAGTTATGTCGATT--GGGTTTAAACATATTTGGAAAG	1138
239	Db		
1139	2y	AATGAGTTGCCAATGTGTTCATGTTTGCACGCCGTACTAAAAAAATTTGGATCTCCTCTAT	1198
299	Db		
1199	2y	GCAATGCTAGACACGAGGAGCATTTGTATGTTTAAATCCAAAGGTGTCCAAATCTGGAAGTC	1258
359	Db		
1259	2y	CTTGAGACAAGGAATGTAATTTGGAGATAGAGGTTTAGAGGTTCTTGGTCTGTGTTGTAAG	1318
419	Db		
1319	2y	AGCTTTAAAGGCTTAGGATTTGAAAGGGCGAGTATGATCAAGGAATCGAGGATGAAGAA	1378
479	Db		
1379	2y	GCTACTGTGTCCATAGAGGGCTAATAGCTTTGTCAACAGGGCTGTTACAGAGCTTGATAC	1438
539	Db		
1439	2y	ATGGCTGTTTATGTGTCGATATTAACAATGCATCTCTGGAACATATTCGGAATCACTCTTG	1498
596	Db		
1499	2y	AGAACTCTGTGAATTTTCGCTTGTGTTTGCATCAAAATGGGGCGCTGGAATCCATTTGGC	1558
656	Db		
1559	2y	CAACTTGACAAATGGGGTGAGGGCTTACTGAGGGGCTGTGACAAGCTGAGGAGATTGCT	1618
716	Db		
1619	2y	CTATATCTCAGCGGTGGCGGTTGACTGATGAGGCTTGGTTACATTCGACAAATACAGT	1678
776	Db		
1679	2y	CMAATGTGAGATGGATCTGCTTGTGTTATGTGGGAGTCTGATGACAGGGCTTTTGGAG	1738
836	Db		
1739	2y	TTTCGCTAAGGGGTGCTCTAGTCTTTCAGAACTTTGAAATGAGAGGGGTGTTATTTTTT	1798
896	Db		
1799	2y	GAAACGTGCACTTGTGTGCGTGCNAACAATTCATCTCTTAGGTACTTGTGGGTGCAA	1858
953	Db		
1859	2y	GTTTATCGGTGTAICTCATCTGACCGTGATCTTTTGTGTAATGGCTCGACCCCTTTTGAAC	1918
1013	Db		
1012	Db	GGGTACAGAGCCTTCGAAAGGGCGCAGGACCTCATGCTCATGGCCAGCCGTACTCGAAC	1072

Tue Apr 20 10:32:12 2004

us-10-009-791-21.rnpb

```

1919 ATTGAGTTGATCTCTTAG---AAAGGTGGCTACGAATACCAATCCAGATGAGACTGTA 1975
1073 ATTGAATTCGCCCTCCAGTCCCGAGAGCGCTTATCGGGTATGATGGCAGATGGACAGCGCT 1132
1976 GTTGTGAGACATCGCTCATATCTTTCGATATATCTTTCGAGGCGAGATCAGAT 2035
1133 TGTGTTGACACTCATGCCAGGTCCTTTCGCTATTAATCTTTCGCTGGAAGGAGCGCGAC 1192
2036 TTTCAGATCTGTTGCTGCTTGGCACTGCCACA 2071
1193 TGCCTCTCAGTGGCTGATGCTTGCACCTTCGCTCA 1228

RESULT 9
US-10-424-599-125710
; Sequence 125710, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 125710
; LENGTH: 669
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(669)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_84525C.1
US-10-424-599-125710

Query Match 14.1%; Score 323; DB 13; Length 669;
Best Local Similarity 81.3%; Pred. No. 2.5e-88;
Matches 374; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 1628 AGCGTGGCGGTTGACTGATGTAGCCCTTGTATGATGGAACAATGATGCAATGATG 1687
DB 1 AGACCTGGGGATGACTGATGTGGCTAGGTACGTAGGCAATACAGCCCAACGTTG 60
QY 1688 AGATGATGCTGCTTGGTTATGTTGGGGAGTCTGATGCAAGGCTTTTGGAGTTGCTAAG 1747
DB 61 AGATGATGCTTCTTGGTTATGTTGGGAGAGACTGATGCAAGGCTTTTGGAAATCTTAAG 120
QY 1748 GGGTGTCTAGTCTTCAGAACTTGAATGAGAGGGTGTATTTTTCAGTGAACGTGCA 1807
DB 121 GGTGCCCCAGTCTTCAGAACTTGAATGAGAGGGGATGCTCTCTTTCAGTGAATGCA 180
QY 1808 CTTGCTGTGGCTGCAACACAAATGACTTCTCTTAGTATCTTGGGTGCAAGTTAGT 1867
DB 181 CTAGCTATTGCTGCACTCACTGAAATCTCTCAGGTACCTATGAGGTGCAAGGTTAGT 240
QY 1868 GTATCTCCATCTGCACTGCTTCTTGGTAAATGCTGCAAGGCTTTTGGCAATTTAGTTG 1927
DB 241 GCATCTGATCTGACGCGATCTTCTGCAATGCTGCAAGGCTTTTGGCAATTTAGTTG 300
QY 1928 ATTCTCTTGAAGAGGTGGCTACGAATACCAATCCAGATGAGACTGTAGTTTGGACAT 1987
DB 301 ATTCTCTTGAAGAGGTGGTTGATGCAATCAGCAAGAGAGCTGTAGTTTGGAGCAC 360
QY 1988 CCGTCTCATATCTTGCATATATTTCTTCTCGAGGCGAGATCAGATTTTCCAGACT 2047
DB 361 CTGACTCATATCTTGCATATATTTCTTCTGCTGCCCAAGACAGATTTTCCAGATACC 420
2048 GTTGTGCTTTGGACACTGCCACATGCGTTGATACCTAGA 2087

```

```

DB 421 GTAATACCTTTGGATCCTGGAAACATATATGTTGCACACCTAAA 460

RESULT 10
US-10-424-599-125711
; Sequence 125711, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 125711
; LENGTH: 744
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(744)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_84526C.1
US-10-424-599-125711

Query Match 13.6%; Score 311.6; DB 13; Length 744;
Best Local Similarity 75.1%; Pred. No. 9.1e-85;
Matches 431; Conservative 0; Mismatches 134; Indels 9; Gaps 3;

QY 1637 GGGTTGACTGATGATGAGGCTTTGTTTACAT-TGGACAATACAG-TCCAAATCTGAGATGGA 1694
DB 2 GGATTGACTGATGATGAGGCTTCGGTTTACATAGGGGCAATACAGCCCCAACGCTGAGATGGA 61
QY 1695 TGCTGCTTGGTTATGTTGGGGAGTCTGATGCAAGGCTTTTGAGTTGCTGAAGGGTCTC 1754
DB 62 TGCTTCTTGGTTATGTTGGGAGAGACTGATGCAAGGCTTTTGGAATCTCTAAGGGTTCCT 121
QY 1755 CTAGCTTTCAGAACTTGAATGAGAGGCTGTTTATTTTTCAGTGAACTGACCTTGGTCT 1814
DB 122 CCAGTCTTTCAGAACTTTCAGATGAGGGGATGTTCTCTTTCAGTGAGTATGCTAGCTA 181
QY 1815 TGCTGCAACAAATGACTTCTCTTAGTACTTCTGCTGCTGCAAGGTTATGCTGATCTC 1874
DB 182 TTGCTGCAACTCACTGAAATCTCTCAGTACTATGGTGCAGGGTATGGAGCATCTC 241
QY 1875 CATCTGGACGCTGATCTTTTGGTAAATGGCTTCGACCCCTTTTGGACATTTGATTCCTT 1934
DB 242 CATCTGGACGCTGATCTTTTGGCAATGGCTTCGCCCTTATTTGGAACTTGAATTCCTT 301
QY 1935 CTAGAAAGGTGGCTAGCAATACCAATCCAGATGAGACTGTAGTTTGGAGCATCTGCTC 1994
DB 302 CTAGACGCTTGTGTTAGCAATCAGAAAGAGAGCCCGTAGTTTGGACACCCCGGCTC 361
QY 1995 ATATCTCTGCATATATTTCTTTCGAGGCGCAGAGATCAGATTTCCAGATCTGTTGTC 2054
DB 362 ATATCTCTGCATATATTTCTTTCGAGGCGCAGAGATTTCCAGATCTGTTTATAC 421
QY 2055 CTTTGGACACTGCCAATGCTGTTGATACCTAGA-----GGCCAGGCTGTGTATAT 2107
DB 422 CTTTGGATCTCTGCAACATATGTTGACACCTAACTTTGTACATACCAATTTTTTTTCTTC 481
QY 2108 ACCAGTTTCTTTTGTGTTTCTTCTCCCTTTCATATGCTGTTTCTATGTTCTGCTCTA 2167
DB 482 TTCTTCTTCCAAAGGTTTTTTTTTTTCTCTTTCATATCTTGTGTTGTAAGTGGTGGAGTG 541
2168 TTTGTAGTTTCATTTTAGACAATTAGTCTTTGTAAT 2201

```

Db 542 TTTTAGTTAGAGGTTTCAATTTCTCTGCAAT 575

## RESULT 11

US-10-424-599-317  
; Sequence 317, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 317  
; LENGTH: 829  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_100286C.1  
US-10-424-599-317

Query Match 12.7%; Score 290.4; DB 13; Length 829;  
Best Local Similarity 74.4%; Pred. No. 3.5e-78;  
Matches 366; Conservative 0; Mismatches 126; Indels 0; Gaps 0;  
  
QY 319 GGAGGAACGGAACGTGGGAGACACGTGTGGTCGACGTGGTCTCGACTCGCTCATCC 378  
Db 245 GGACGGAGCGGATGACGGGAGGATTATCGGATGGTCTGGACTGGTGAATGCC 304  
QY 379 TTATCATGACGACCCCAAGACCGGACCGCGTTTCCAGGTGTGTCAGCGTGGTACGA 438  
Db 305 ATACATACACGACTCCAAAGGACCGGACCGGTGTGCAGGTGTGCAGGCGGTGGTACGA 364  
QY 439 GCTCGACTGCTCACCGGCAAGACGTCACATGCGGCTCTGTACACACACCGCGGC 498  
Db 365 GCTCGACTGCTCACTCGTAAACACGTCACATTTGGGCTGTGTATACACACCGCGCGGA 424  
QY 499 TCGCTCCGCGCGCGCTTCCGACCTCGAGTCGCTCAAGCTCAAGGCAAGCCCGAGC 558  
Db 425 CCGGCTTAAGCGCGGTTCCGACCTCGAGTGTCTGAAGTGTGAGGAAAGCGCGGC 484  
QY 559 CGCAATGTTCAACTTGATACCCGAGGATGGGCGGACACGTCACTCCCTGGGTCAAAGA 618  
Db 485 GCGGATGTTCAACCTGATACCTGAGGATTGGGAGGTTTGTGTACGCGCTTGGGTCAAGT 544  
QY 619 GATTTCTAGTACTTGTATGCTCTCAAGAGCTCCACTTCGCGCGCATGATTTGTCAGGA 678  
Db 545 GATCTCTCAATCTTTGATTTGTTGAAGTCGCTTCACTTTTCGCGCGCATGATTTGTGAGGA 604  
QY 679 TTCGATCTTTCAGAACTCTCGCTCGTGAACCGCGGTCAAGTCTTCAAGCTTGA 738  
Db 605 TTGCGATCTTCAAGTCTACTGTTTCGCGCGGTCAAGTCTTCAAGCACTCAGGCTGGA 664  
QY 739 CAAGTGTCCGGTTTCAACCGCATGCTTTTCCATATCGGTGCTTTCGAAAGGTTT 798  
Db 665 AAATTTGCTGGATTCTCCACTGATGCGCTCTACTATAGTCAITTCGAATTAAGTGTTA 724  
QY 799 AAGAGTCTTGT 810  
Db 725 GGCTCACTTGT 736

## RESULT 12

US-10-424-599-44612  
; Sequence 44612, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 44612  
; LENGTH: 428  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_140283C.1  
US-10-424-599-44612

Query Match 12.2%; Score 279.8; DB 13; Length 428;  
Best Local Similarity 80.9%; Pred. No. 4e-75;  
Matches 326; Conservative 0; Mismatches 77; Indels 0; Gaps 0;  
  
QY 577 ACCGAGGATTGGGCGGACACGTCTACTCCCTGGGTCAAAGAGATTTCTCAGTACTCGA 636  
Db 25 ACCGAGGATTGGGAGGTTTGTCTACGCGGGGGTCCAGAGAGATCTCTCAGTACTCGA 84  
QY 537 TTGCCTCAAGAGCTCCACTTCGCGCGCATGATGTCAAGGATTCGATCTCAGATCT 696  
Db 85 TTGCTTGAAGTCGCTTCACTTTTCGGGTATGATTGTAAGGATTCGGAATCTTCAAGTCT 144  
QY 697 CGTCTCGTACCGCGGTCAAGTCTTCAAGCTTCAAGCTTGACAAGTGTCCGTTTCAC 756  
Db 145 AGTCTGTCGCGGACACATCTTCAAGCACTCAAGCTGGAAGTGTCTGATTTCTC 204  
QY 757 CACGATGCTTTTCCATATGCTGCTTTTGCAGAGTTTAAGAGTCTTGTTTTGA 816  
Db 205 CACGATGCTCTACTATATGCTGCTTTATTCAGGAACTTAAGAGTCTTGTTTTGA 264  
QY 817 GGAAGCTCAATCTTGAAGACCGAGATGGCTACAGAGCTTGTCTTGAATATAC 876  
Db 265 GGAAGCTCACTTGTGGAATGATGTGATGCTGCTCCATGAGCTTGTCTTGAATATAC 324  
QY 877 AGTCTTGAAGTCTCAATTTTCTTGAAGCAATGCTTGTGGAAGATTCAGGACCT 936  
Db 325 AGTCTTGAAGCACTGAATTTTCTTGAATATGCTTGTGGAAGATTCAGGACCT 384  
QY 937 TGAATCTTGAATATAAATTTGCCCAACTTGAATGCTGTGAAA 979  
Db 385 TGAATCTTGAAGCAATTTGCCCAACTTGAATGCTGTGAAA 427

## RESULT 13

US-09-770-149-161  
; Sequence 161, Application US/09770149  
; Patent No. US20020059663A1  
; GENERAL INFORMATION:  
; APPLICANT: Gorlach, Jorn  
; APPLICANT: An, Yong-Qiang  
; APPLICANT: Hamilton, Carol M.  
; APPLICANT: Price, Jennifer L.  
; APPLICANT: Raines, Tracy M.  
; APPLICANT: Yu, Yang  
; APPLICANT: Rameaka, Joshua G.  
; APPLICANT: Page, Amy  
; APPLICANT: Matthew, Abraham V.  
; APPLICANT: Ledford, Brooke L.  
; APPLICANT: Woessner, Jeffrey P.  
; APPLICANT: Haas, William David  
; APPLICANT: Garcia, Carlos A.  
; APPLICANT: Krickler, Maya  
; APPLICANT: Slader, Ted  
; APPLICANT: Davis, Keith R.  
; APPLICANT: Allen, Keith  
; APPLICANT: Hoffman, Neil



; APPLICANT: Hurban, Patrick  
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis  
; TITLE OF INVENTION: thaliana  
; FILE REFERENCE: 2024 (PARA-013PRV)  
; CURRENT APPLICATION NUMBER: US/09/1770,149  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/178,506  
; PRIOR FILING DATE: 2000-01-27  
; NUMBER OF SEQ ID NOS: 999  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 161  
; LENGTH: 719  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; ORGANISM: Arabidopsis thaliana  
; US-09-770-149-161

Query Match 9.9%; Score 227.2; DB 9; Length 719;  
Best Local Similarity 66.7%; Pred. No. 1.1e-58;  
Matches 356; Conservative 0; Mismatches 173; Indels 5; Gaps 2;

QY	1593	GCTGTGACAGCTGAGGAGATTGGCTCTATCTCAGGCGTGC	CGGTTGACTGTAGTAG 1652
DB	11	GATGCAAGAACCTCAGACGATTTCGATTTCTATCTGAGACA	GGCGCTTAACTCGACTTGG 70
QY	1653	GCCTTGGTTACATTGGCAATACAGTCCAAATGTGAGATCG	ATGCTTGTGTTATGTGG 1712
DB	71	GCTTAAGCTACTCGGACACTACAGTCCAAACGTGAGATGG	ATGCTTGGCTTGGCTTACGTAG 130
QY	1713	GGGAGTCTGATCAGGGCTTTGGAGTTCGCTAAGGGGTGCT	AGTCTTCTCGAAACTTGG 1772
DB	131	GTGAATCAGATGAAGGTTTAATCGAATTCCTCAAGAGCG	CTGTCCAAATCTACAGAAGCTAG 190
QY	1773	AAATGAGAGGGTGTATTTATTTTTCAGTGAACGCTGCAC	TGTGTGGCTGCACAACAACTTGA 1832
DB	191	AGATGAGAGGTTGT--TGCTTCAGTGAGCGAGCAATCG	CTGCAGCGGTTACAAAATTGC 247
QY	1833	CTTCTCTTAGGTACTTGTGGGTCAAGGTTATGGTGTATC	CTCATCTGACGTGATCTTTT 1892
DB	248	CTTCACTGAGATACTTGTGGGTCAAGGTTACAGAGCATC	GATCGATCGCGGACAAAGATCTAA 307
QY	1893	TGGTAATGGCTCGACCCCTTTTGGCACTTCAGTTGATTC	CTCTAGAAAGGTGGCTTACGA 1952
DB	308	TGCAGATGGCTTAGACCCGTACTGGAACATCAGAGCTG	ATCCATCAAGAAGNGTCCCGAAG 367
QY	1953	ATACCAATCCAGATGAGACTGTAGTTGTGGAGCATCTG	CTCATATTTCTTGCATATTTT 2012
DB	368	TGAATCAACAAGGAGAGATAAGAGAGATGGAGCATCG	CGGCTCATATATTGGCTTACTACT 427
QY	2013	CTCTTGCAGGCGCAGAGATCAGATTTCCAGATACTGTG	TGTCCTTTGGACACTGCCACAT 2072
DB	428	CTCTGGCTGGCCAGAGAACAGATTGTCCAACTGTTAG	AGTCTTGAAGAGGCCAATAT 487
QY	2073	CGCTTGTATACCTAGAGGCCAGAGCTGTGTATATATAC	CAGTTTTTCTTTTGTGTTTT 2126
DB	488	GATATG--ACCCAAAACAGGTTTGTATATAAAGATTT	TTTAGTCTCGAGTTTTT 539

RESULT 14  
US-10-424-599-21331  
; Sequence 21331, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kowalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 21331

```

; LENGTH: 449
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_119267C.1
US-10-424-599-21331

Query Match          9.6%; Score 220.2; DB 13;
Best Local Similarity 94.6%; Pred. No. 1.1e-56;
Matches 228; Conservative 0; Mismatches 13;

Qy 552 CCGAGCGCGAATGTTCAACTTGATACCCGAGGATTTGGGG
Db 138 CTCAAGCGCGAATGTTCAAGTTCGATTCGCGAGGATTTGGGG
Qy 612 TCAAGAGATTTCTCAGTACTTTCGATTCGCTCAAGAGCCT
Db 198 TCAAGAGATTTCTCAGTACTTTCGATTCGCTCAAGAGCCT
Qy 672 TCAAGGATTCGCGATCTTCAGAAATCTCGCTCGTACCGCGGG
Db 258 TCAAGGATTCGCGATCTTCAGAAATCTCGCTCGTACCGTGGG
Qy 732 AGCTTGACAAGTCTCCGGTTTCACACCGATGTCCTCTTTT
Db 318 AGCTTGACAAGTCTTCAGTTCAGTTCACCACTGATGCTCTTTT
Qy 792 A 792
Db 378 A 378

RESULT 15
US-10-424-599-21332
; Sequence 21332, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and
; TITLE OF INVENTION: plants and Uses Thereof for Pl
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 21332
; LENGTH: 789
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_119268C.1
US-10-424-599-21332

```

Query Match	9.6%;	Score 218.6;	DB 13;	Length 789;
Best Local Similarity	94.2%;	Prod. No. 5.2e-56;		
Matches 227;	Conservative	0;	Mismatches 14;	Indels 0; Gaps 0;
552	CCCGAGCCGAATGTTCAACTTGATACCGGAGGATTGGGCGGACACGTCATCTCCCTGGG	611		
127	CTCAAGCCGAATGTTCACTTGATTCCGAGGATTGGGCGGAACATGTTAGTCCCTGGG	186		
612	TCAACAGAGATTTCCTCAGTACTTCGATTGCCTCAAGAGCCTCCACTTCGGCGCATGATTG	671		
187	TCAACAGAGATTTCCTCAGTACTTCGATTGCCTCAAGAGCCTCCACTTCGGCGCATGATTG	246		
672	TCAAGGATTCCGATCTTCAGAAATCTCGTCTGTGACCGCGGTCAAGTGTCTCAAGCTCTCA	731		
247	TCAAGGATTCCGATCTTCAGAAATCTCGTCTGTGACCGGTCAAGTGTCTCAAGCTCTCA	306		
732	AGCTTCACAAGTGTCTCCGGTTTCACACACGATGGTCTTTTCCATATCGGTCCGTTTTGCA	791		

us-10-009-791-21.rnpb

Tue Apr 20 10:32:12 2004

Db 307 AGCTTGACAGTGCTTCAGTTTCACCACTGANGGTCCTTTCCACATCGGTCGCTTTTGCA 366

Qy 792 A 792

Db 367 A 367

Search completed: April 20, 2004, 01:34:07  
Job time : 1495 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 18:00:14 ; Search time 5811 Seconds  
(without alignments)  
11757.816 Million cell updates/sec

Title: US-10-009-791-21  
Perfect score: 2288  
Sequence: 1 gcacgagccacacgttaca.....acattcgaattttgtgaat 2288

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: em\_estba:\*  
2: em\_estba:\*  
3: em\_estba:\*  
4: em\_estba:\*  
5: em\_estba:\*  
6: em\_estba:\*  
7: em\_estba:\*  
8: em\_estba:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_est3:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estba:\*  
16: em\_estba:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pin:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_pbg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	712	31.1	741	10	BE821935
2	614.2	26.8	2298	11	AY109490
3	598.2	26.1	2190	11	AY109565
4	564	24.7	588	14	CA820059

5	555.8	24.3	1811	14	CA654430
6	543.4	23.8	560	14	CF807278
7	542	23.7	574	13	BQ297870
8	538	23.5	538	13	BQ133928
9	532.2	23.3	561	12	BI967384
10	526	23.0	541	14	CF807291
11	515.6	22.5	538	12	BI974579
12	493.4	21.6	550	14	CA820349
13	489	21.4	493	9	AI855554
14	476.2	20.8	555	13	BQ630213
15	464	20.3	574	10	AW759107
16	462.6	20.2	589	10	BF325270
17	460.8	20.1	658	12	EG453637
18	451.6	19.7	783	14	CB894250
19	435.8	19.0	441	10	BE802940
20	417.8	18.3	455	12	BG156552
21	413.2	18.1	704	14	CB008682
22	391.2	17.1	630	14	CB001376
23	389.4	17.0	693	13	BU894657
24	385.6	16.9	465	10	AW396651
25	377.2	16.5	670	13	BU867812
26	375.2	16.4	542	13	BU080606
27	374.2	16.4	596	14	CB919625
28	373.8	16.3	512	10	AW185771
29	372.8	16.3	592	14	CB920883
30	372.2	16.3	442	10	BG045553
31	368	16.1	1683	11	AY109486
32	363.2	15.9	739	13	BQ997606
33	359.8	15.7	790	29	CG964083
34	359.6	15.7	382	10	AW595938
35	351.6	15.4	1134	14	CK209957
36	351.4	15.4	563	14	CB923093
37	351.4	15.4	813	28	BH678933
38	348.8	15.2	570	14	CB911803
39	341.4	14.9	556	12	BI892238
40	333.8	14.6	950	29	CG945532
41	333	14.6	652	13	BU548376
42	331	14.5	479	10	BG044386
43	330	14.4	428	14	CA936527
44	329.4	14.4	586	9	AI482980
45	326.8	14.3	588	10	AW222912

ALIGNMENTS

RESULT 1

BE821935/c  
LOCUS  
DEFINITION

BE821935 741 bp mRNA linear EST 24-MAY-2001  
GM700015B20H12 Gm-r1070 Glycine max cDNA clone Gm-r1070-6144 3',  
mRNA sequence.

ACCESSION BE821935

VERSION BE821935.1

KEYWORDS EST

SOURCE Glycine max (soybean)

ORGANISM Glycine max

REFERENCE 1 (bases 1 to 741)

AUTHORS Vokkin, L., Kelm, P., Shoemaker, R., Retzel, E., Khanna, A., Corvett, V.,

Epelting, J., Raph, C., Shoop, E., Pardinas, J., Liu, L. and Lewin, H.

TITLE A Functional Genomics Program for Soybean (NSF 9872565)

JOURNAL Unpublished (1999)

COMMENT Other\_ESTs: AW185771 corresponding to Gm-cl019-1292 (5')

Contact: Vokkin, L.O., PI, A Functional Genomics Program for

Soybean (NSF 9872565)

Lewin, H. A., Director, Keck Center for Comparative and Functional

Genomics

University of Illinois

Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA

Tel: (217) 244-6147

Fax: (217) 333-4582  
 Email: 1-vokine@uic.edu  
 This clone is available through: Genome Systems, Inc. 4633 World  
 Parkway Circle St. Louis, Missouri 63134. For further information  
 call: (800) 430-0030 or (314) 427-3222 FAX: (888) 913-3324 or (314)  
 427-3324 or contact: clones@genomesystems.com or info@genome  
 systems.com web site: www.genomesystems.com  
 Seq primer: 5'-TTTTTTTTTTTTTTTTT (A/C/G)-3'.

## FEATURES

source

1. 741  
 /location/Qualifiers  
 /organism="Glycine max"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3847"  
 /clone="Gm-r1070-6144"  
 /clone\_lib="Gm-r1070"  
 /note="The library Gm-r1070 is a sequence-driven, reracked  
 set of 9,216 clones selected from cDNA libraries from  
 various tissues and stages of development of soybean that  
 represent 2,639 sequences from immature cotyledons, 1,770  
 from immature seed coats, 3,938 from flowers, and 869  
 from young pods. The 5' ESTs of the source clones from  
 the different libraries was used to select singletons, or  
 a representative of each contig, which were reracked to  
 form library Gm-r1070. The cDNA clones of the reracked  
 Gm-r1070 library were then sequenced at the 3' end. The  
 contig analysis to select unique genes was performed by  
 the laboratory of Ernest Retzel, Center for Computational  
 Genomics and Bioinformatics, University of Minnesota.  
 http://www.cbc.umn.edu/ResearchProjects/Soybean/index.html  
 . Reracking was performed by Genome Systems, St. Louis,  
 http://www.genomesystems.com, and 3' sequencing by the  
 Keck Center for Comparative and Functional Genomics,  
 University of Illinois.  
 http://www.life.uiuc.edu/biotech/keck.html. Note: The  
 corresponding 5' EST from each clone in the Gm-r1070  
 library is listed in the 'OTHER EST' field. The detailed  
 information on the source library for each clone can also  
 be obtained by referring to the Genome Systems clone ID of  
 the original cDNA library that is also listed under  
 'OTHER EST'."

## ORIGIN

Query Match 31.1%; Score 712; DB 10; Length 741;  
 Best Local Similarity 96.1%; Pred. No. 7.2e-123;  
 Matches 712; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 1523 GTGTTGTTGACCATGAAGAAGATACATGTTGCCACTTGACATGGGGTGGGCT 1582  
 DB 741 GTGNGCNCNNNCCATGAAGAAGATNCTGATTTGCCACTNGNNAATGGGTGAGGNN 682

QY 1583 CTACTGAGGGCTGTGCAAGCTGAGGAGATTGCTCTATATCTCAGCGTGGCGGGTTG 1642  
 DB 681 CTACTGAGGGCTGTGCAAGCTGAGGAGATTGNNCTATANNCCAGCGTNNCGGGTTG 622

QY 1643 ACTGATGTAGGCTTGGTTTACATTGCGACATACAGTCCAAATGTGAGATGATGCTT 1702  
 DB 621 ACTGATGTANNNNNTGGTTTACATTGNACATACAGTCCAAATGTGAGATGATGCTT 562

QY 1703 GGTATGTGGGGAGTCTGATGACAGGCTTTGGAGTTGCTTAAGGGGTGCTAGTCTT 1762  
 DB 561 GGTATGTGGGGAGTCTGATGACAGGCTTTGGAGTTGCTTAAGGGGTGCTAGTCTT 502

QY 1763 CAGAACTTGAATGAGGGGTGTTTATTTTTCAGTGAACGTGCATCTGCTGGCTGCA 1822  
 DB 501 CAGAACTTGAATGAGGGGTGTTTATTTTTCAGTGAACGTGCATCTGCTGGCTGCA 442

QY 1823 ACACAAATGACTTCTTCTAGGTACTGTGGGTGCAAGGTATGTTGTTATCTCCATCTGA 1882  
 DB 441 ACACAAATGACTTCTTCTAGGTACTGTGGGTGCAAGGTATGTTGTTATCTCCATCTGA 382

QY 1883 CGTGATCTTTGGTAAATGCTCGACCCCTTTTGGACATTCAGTTGATCTTCTAGAAAG 1942  
 DB 381 CGTGATCTTTGGTAAATGCTCGACCCCTTTTGGACATTCAGTTGATCTTCTAGAAAG 322

QY 1943 GTGGCTACGAATACCAATCCAGATGAGACTGTAGTTGTTGAGCATCTGCTCATATCTT 2002  
 DB 321 GTGGCTACGAATACCAATCCAGATGAGACTGTAGTTGTTGAGCATCTGCTCATATCTT 262

QY 2003 GCATATATTCTCTTCGAGGGCAGAGATCAGATTTTCCAGATCTGTTGTCCTTTGGAC 2062  
 DB 261 GCATATATTCTCTTCGAGGGCAGAGATCAGATTTTCCAGATCTGTTGTCCTTTGGAC 202

QY 2063 ACTGCCACATCGTTGATACCTAGAGCCAGAGCTGTGTATATATACACGTTTCTTTTG 2122  
 DB 201 ACTGCCACATCGTTGATACCTAGAGCCAGAGCTGTGTATATATACACGTTTCTTTTG 142

QY 2123 TTTTCTCTCCCTTTTCATATGCTGTTCATGTTCCTGCTCTATTTAGTTCATTT 2182  
 DB 141 TTTTCTCTCCCTTTTCATATGCTGTTCATGTTCCTGCTCTATTTAGTTCATTT 82

QY 2183 AGACAATTAGTCTGTAAATAGCCTGTGTTTTCATTTGAAATCTGAAAGCTTCCCTTA 2242  
 DB 81 AGACAATTAGTCTGTAAATAGCCTGTGTTTTCATTTGAAATCTGAAAGCTTCCCTTA 22

QY 2243 ACCGTATTGGCTCCCTTAAAA 2263

DB 21 ACCGTATTGGCTCCCTTAAAA 1

## RESULT 2

AV109490  
 LOCUS AV109490 2298 bp mRNA linear HTC 17-OCT-2002  
 DEFINITION Zea mays CL920\_1 mRNA sequence.  
 ACCESSION AY109490  
 VERSION AY109490.1 GI:21213230  
 KEYWORDS HTC.

## SOURCE

Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.  
 1 (bases 1 to 2298)  
 Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,  
 Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.  
 Maize Mapping Project/DuPont Consensus Sequences for Design of  
 Overgo Probes  
 Unpublished (2002)  
 2 (bases 1 to 2298)  
 Coe, E.H.

## REFERENCE

1 (bases 1 to 2298)  
 Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,  
 Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.  
 Maize Mapping Project/DuPont Consensus Sequences for Design of  
 Overgo Probes  
 Unpublished (2002)  
 2 (bases 1 to 2298)  
 Coe, E.H.

## AUTHORS

Maize Mapping Project/DuPont Consensus Sequences for Design of  
 Overgo Probes  
 Unpublished (2002)  
 2 (bases 1 to 2298)  
 Coe, E.H.

## TITLE

Direct Submission  
 Submitted (25-APR-2002) Maize Mapping Project, University of  
 Missouri, Columbia, MO 65211, USA

## JOURNAL

If you are interested in getting corresponding physical clones,  
 these are publicly available from ZmDB and may be found by BLAST  
 searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,  
 www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the  
 maize cDNA sequences is either Virginia Walbot, Stanford or Pat  
 Schnable, Iowa State, then clones may be requested from ZmDB:  
 www.zmdb.iastate.edu.

## FEATURES

Location/Qualifiers  
 1. 2298  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:633078"  
 /db\_xref="taxon:4577"  
 /clone\_lib="Maize Mapping Project/DuPont Consensus  
 Library"  
 /note="this sequence is part of a project of EST  
 assemblies resulting from the application of public  
 contigs to seed DuPont contigs; this resource was  
 assembled by DuPont as part of a collaboration for the  
 overgo addressing of BACs in conjunction with the Maize  
 Mapping Project"

## ORIGIN

Query Match 26.8%; Score 614.2; DB 11; Length 2298;



Tue Apr 20 10:32:13 2004

www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the  
maize cDNA sequences is either Virginia Walbot, Stanford or Pat  
Schnable, Iowa State, then clones may be requested from ZmDB:  
www.zmdb.iastate.edu.

# FEATURES

Location/Qualifiers  
1..2190  
/organism="Zea mays"  
/mol\_type="mRNA"  
/db\_xref="MaizeDB:633027"  
/clone\_lib="Maize Mapping Project/DuPont Consensus  
Library"  
/note="this sequence is part of a project of EST  
assemblies resulting from the application of public  
contigs to seed DuPont contigs; this resource was  
assembled by DuPont as part of a collaboration for the  
overgo addressing of BACs in conjunction with the Maize  
Mapping Project"

## ORIGIN

Query Match 26.1%; Score 598.2; DB 11; Length 2190;  
Best Local Similarity 60.6%; Pred. No. 1.3e-101;  
Matches 1053; Conservative 0; Mismatches 665; Indels 21; Gaps 6;  
345 GTGTGGTGCAGTGGTCTCGACTGGCTCATCCCTTACATCGACGACCCCAAGGACCGG 404  
231 GCCTCCCGAGGAGGGCTGCACCTGTGTTCGGCTACATGGACGACCCCGGAGACCGG 290  
405 ACGCCGCTTTCCAGGTGTGTCAGCGTGTGACGATCGACTCGCTCACCGGCAAGCAG 464  
291 AGCGCGCTCGCTCGTGTGCGGCTCTGGCACCGCATCGACGGCTCTCGCGAAGCAG 350  
465 TCACCATGGCGCTCTGCTACACACACCGCGCTCGCTCCGCGCGCTTCCCGCAC 524  
351 TCACCGTGGCGCTTCTGCTACCGCGTGGAGCCGCGCGCTGCTCGCAGGTTCCGCGGC 410  
525 TCAGTGGCTCAAGTCAAGGCAAGCCCGGAGCGGCAATGTTCAATTCGATACCGGAG 584  
411 TCAGTGGCTCGCGCTCAAGGAGGCGCCGCGCGCATGTACGGGCTCATCCCGAAG 470  
585 ATTGGGCGGACACGCTCACTCCCTGGGTCAAGAGATTTCTCAGTACTTCGATGGCTCA 644  
471 ACTTTGGCGCTACGCGCGCTCGGTTCGCGGAGCTCGCGCGCTTCTGCTGCTCA 530  
645 AGAGCTTCGCTTCGCGCGCATGATTTCAAGGATTCGATCTTCAGATCTCCTCGTG 704  
531 AGGCGCTCCACTCGCTGCGATGACCGTTCACGAGGACATCGCGTGTCTCTCCAG 590  
705 ACCGCGTCAAGTCTTCAAGCTCTCAAGCTTGAAGTGTCTCGGTTTCAACCGAGT 764  
591 CGCGCGGCAATGCTACAGCGCTCAAGCTCGACAGTGTCTCGGCTTCTCAACAGNN 650  
765 GTCTTTTCCATATCGTCTGCTTTGCAAGATTTAAAGATCTTGTTTTGGAGAAAGCT 824  
651 NNN 710  
825 CAATTTCTTGAAGAGACGAGATGCTGCTACAGGCTTGTCTTGAATATACAGTTCTTG 884  
711 TAAATGAGATGAAGGAGTGAATGGCTCCATGAATCGCTGTCAACATTTCTGTTCTGG 770  
885 AGACTCTCAATTTTACITGACAGACATTTGCTGTGTGAAGATTTGAGACCTTTGAATTT 944  
771 TGACACTGAATTTTACATGACAGAACT--CAAGTGGAGCGCTGCTGATCTGGAGCTTC 827  
945 TAGCTAAATAATGGCCCAACTTAGTGTCTGTAAGACTTACTGACTGTGAATTTACTGGATC 1004  
828 TTGCAAGAAGACTGTAATCAITTAATTTCTTTGAAGATGGGTGACTGTGATCTTTCAGATC 887  
1005 TTGTGAACTCTTTAAGCATGCTCTGCGCTGGAAGAGTTTGTGGAGGCACTTCAACG 1064  
888 TGATTGGGNNNNNNCCAAACCTCAAGAGATTTGCAAGATTTCCGCGGAGTGCNNNNNG 947  
1065 AGGAACACGA-----AAGATACTCTGCTATATCATTACCAGCAAGTTATGTGATT-- 1116

Db 948 AAGTTGAGAGTACACAGTACGAAAGGTCATTTTCCACCTAGGTCATGCTTCTTAN 1007  
QY 1117 -GGGTTTAAACATATATTGGAAGAATAGTTGGCCATTTGTTTTCATGTTTCAGCGGTAC 1175  
Db 1008 NNNNTCTTACCTTCATGGTNNNNNCGAGATGCCAGTTATCTTTCCGTTATTCACAAATGC 1067  
QY 1176 TAAAAAATTTGGATCTCTCTATGCAATGCTAGACAGGAGATCATGTATGTTAAATCC 1235  
Db 1068 TTAAGAAATCTGGACTTGCAGTTCTTCCCTCACCACTGAAGACCACTGTGAGCTTATTG 1127  
QY 1236 AAGGTGTCCAAATCTGGAATCTTGTGAGACAAGGAATGTAATTTGAGATAGAGGTTAG 1295  
Db 1128 CTAATGCCCCCAACCTAAGTGTCTTCGAGGTGAGGATGTGATTGGGATAGAGACTAG 1187  
QY 1296 AGTTCTTGGTGTGTTTGTAAAGGCTTAAAGGCTTAGGATTTGAAAGGGGCGGATGATG 1355  
Db 1188 AAGTTGTTGCTGTACTACATGCAAGAGCTACGAAGGCTCAGAATTTGAGGCTGGCAGATG 1247  
QY 1356 ATCAAGGAATGGAGATGAAGAGGTACTGTGTCCCATAGAGGCTAATAGCTTTGTCTAC 1415  
Db 1248 ATCTCGGCCCAAGAGAGAGAGGAGG--AGTCTCTCAGATAGTTTGACGCGTAGCGG 1304  
QY 1416 AGGCTGTTCAGAGCTTGAATAACATGCTGTTTATGTCTGATATTAACAATGATCTC 1475  
Db 1305 TTGGCTTCCCGCAACTAGAGTACATAGCTGCTGCTATGATCTGATATCAAAATGGGCGC 1364  
QY 1476 TCGAACAATTTGGAACTCACTTGAAGAACCTCTGTGATTTTCGCTTGTGTTGCTTGAC 1535  
Db 1365 TGGAAATCCATTGGCACTTCTGCAAGATCTATATGATTTCCGCTGTGTTCTACTTGACA 1424  
QY 1536 ATGAAGAGAGATTAAGTATTTGCCACTTGAACAATGGGCTGAGGGCTCTACTGAGGGCT 1595  
Db 1425 AACAGAAAGAGATAACAGACCTGCCACTTGAACAAGCGGTCCGTGCCCTACTCAGGA 1484  
QY 1596 GTCAAGCTGAGGAGATTTGCTCTATATCTCAAGCGTGGCGGGTGAAGTGTAGGCGC 1655  
Db 1485 GCGTCAACTTCGAGGTTTGCCTTCTACTGAGACGGGAGGGCTTTCAGACGTAGGCG 1544  
QY 1656 TTGGTTTACATTTGACAATACAGTCCAAATGTGAGATGATGCTGCTTGTGTTATGTTGG 1715  
Db 1545 TCGTTTACATTTGACTATACAGTGGGAACATCCAAATACATGCTGCTCGGCAAGCTTGGT 1604  
QY 1716 AGTCTGATGACGGCTTTGGAGTTCGTAAGGGGTGCTCTAGTCTTTCAGAAACTTGAAA 1775  
Db 1605 AATCTGACAAAGGATTAACAGTTTGAACGGGATGCAACCACTTCGCGAAGCTCGAGC 1664  
QY 1776 TGAGAGGCTGTTTATTTTTCAGTGAACGTCACCTTGTGCTGGTGGAACCAATGACTT 1835  
Db 1665 TGAGGGGCTG--CTGCTTCAGCGAGCGAGCTCTGGCAGTGGCGCTGCTCCAGATGCCCT 1721  
QY 1836 CTCTTAGGTACTTGTGGGTGCAAGGTTATGCTGATCTCCATCTGGACGCTGATCTTTGG 1895  
Db 1722 CGCTGAGTACATATGGGTGCGAGGGTACAGAGCCTCTCGAAACGGGCGAGGACCTCATGC 1781  
QY 1896 TAATGGCTCCAGCCCTTTTGGAACTTTCAGTTT---GATTCTCTTCTAGAAAGTGGCTACGA 1952  
Db 1782 TCATGGCCAGCCGCTACTGGAACTTGAATTCGCGCTCCCATTCCTCCGAGAGCGCTTATC 1841  
QY 1953 ATACCAATCCAGATGAGACTGTAGTTGTGAGCATCTCTGCTCATATTTCTTGCATATATT 2012  
Db 1842 GGGTGTGGCAGATGGACAGCCTTGTGTGACACCCATGCCAGGCTCTTCCGTTACT 1901  
QY 2013 CTCTTGAGGGCAGAGATCAGATTTTTCAGATCTGTTGCTGCTTGGACACTGCCACA 2071  
Db 1902 CCCTTGTGGAAGAGGCGGACCTGCCCTCAGTGGTGTGACCTTTCACCCCTGGCTCA 1960

## RESULT 4

CA820059

LOCUS

DEFINITION

CA820059 588 bp mRNA linear EST 09-DEC-2002  
sau83hl1.1 y1 Gm-cl048 Glycine max cDNA clone SOBEAN CLONE ID:  
Gm-cl048-5685 5' similar to TR:004197 004197 HYPOTHETICAL PROTEIN.



```

; mRNA sequence.
ACCESSION CA820059
VERSION CA820059.1 GI:26268996
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max (soybean)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 588)
Shoemaker,R., Keim,P., Vodkin,L., Epelding,J., Coryell,V.,
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntville, AL 35801 For further information
call: (800)-533-4363 or contact: ccu@resgen.com web site:
www.resgen.com
Seq primer: -40RP from Gibco
High quality sequence stop: 421.
FEATURES
source
1..588
Location/Qualifiers
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl048-2685"
/tissue_type="whole seedlings of greenhouse grown plants"
/dev_stage="1 week old"
/lab_host="DH10B"
/clone_lib="Gm-cl048"
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The Clark NIL was constructed and seed was provided
by Dr. J. Specht, University of Nebraska (Shoemaker and
Specht, 1995). The cDNA library was constructed from mRNA
isolated from whole seedlings of 1 week old greenhouse
grown plants. Complementary DNA was synthesized from mRNA
using a primer consisting of a poly(dT) sequence with a
XhoI restriction site and a 3' anchor. EcoRI adapters were
ligated to the blunt-ended cDNA fragments followed by XhoI
digestion. The cDNA fragments were directionally cloned
into the EcoRI-XhoI restriction site of the pBluescript
vector. The ligated cDNA fragments were transformed into
DH10B host cells (GibcoBRL). The library was constructed
in cooperation with Dr. Paul Keim's laboratory at Northern
Arizona University."

ORIGIN
Query Match 24.7%; Score 564; DB 14; Length 588;
Best Local Similarity 97.4%; Pred. No. 3.4e-95;
Matches 573; Conservative 0; Mismatch 15; Indels 0; Gaps 0;

QY 924 AGATTGAGGACCTTTAGCTTAAATAATGCCCCCACTTAGTGTCTGTGAACCTTA 983
DB 1 AGATTGAGGACCTTTAGCTTAAATAATGCCCCCACTTAGTGTCTGTGAACCTTA 60

QY 984 CTGACTGTGAATACTGGATCTTGTGAACCTTCTTTAGCATGCTCTGGCTGGAGAGT 1043
DB 61 CTGACAGTGAATACTGGATCTTGTGAACCTTCTTTAGCATGCTCTGGCTGGAGAGT 120

QY 1044 TTTGTGAGGACCTTACACGAGAACCCAGAAAGATCTCTGTATATCATTTACCAGCAA 1103
|||||

```

121 TTTGTGAGGACCTTACACGAGAACCCAGAAAGATCTCTGTATATCATTTACCAGCAA 180

QY 1104 AGTTATGTCGATTGGGTTTAAACATATATTGAAAGATGAGTGGCCATTTGTTTCATCT 1163

DB 181 AGTTATGTCGATTGGGTTTAAACATATATTGAAAGATGAGTGGCCATTTGTTTCATCT 240

QY 1164 TTGCAGCGGTACTAAAAAATTGGATCTCTCTATGCAATGCTAGACACGAGGATCATTT 1223

DB 241 TTGCAGCGGTACTAAAAAATTGGATCTCTCTATGCAATGCTAGACACGAGGATCATTT 300

QY 1224 GTATGTTAATCAAAAGGTGTCCAAATCTGGAAGTCTTTGAGACAAAGGAATTAATTGGAG 1283

DB 301 GCATGTTAATCAAAAGGTGTCCAAATCTGGAAGTCTTTGAGACAAAGGAATTAATTGGAG 360

QY 1284 ATAGAGGGTTAGAGGTTCTTGGTCTGTTGTTGTAAGAGGCTTAAAGGCTTAGGATTGAAA 1343

DB 361 ACAGAGGGTTAGAGGTTCTTGGTCTGTTGTTGTAAGAGGCTTAAAGGCTTAGGATTGAAA 420

QY 1344 GGGCGATGATGATCAAGGAATGAGGATGAAGAGGTACTGTGTCCTCATAGAGGGCTAA 1403

DB 421 GGGGATGATGATCAAGGAATGAGGATGAAGAGGTACTGTGTCCTCATAGAGGGCTAA 480

QY 1404 TAGCCTTGTCAAGGGCTGTTTCAGAGCTTGAATACATACATGGCTGTTATGTGTTGATATTA 1463

DB 481 TCGCCTTGTCAAGGGCTGTTTCAGAGCTTGAATACATACATGGCTGTTATGTGTTGATATTA 540

QY 1464 CAATGTCATCTTGAACATATTTGAACTCACTTGAAGACCTCTGTG 1511

DB 541 CAAATGTCATCTTGAACATATTTGAACTCACTTGAAGACCTCTGTG 588

RESULT 5

CA654430 1811 bp mRNA linear EST 25-JUN-2003

LOCUS wreln.pk0122.d3:5p5 wreln Triticum aestivum cDNA clone

DEFINITION wreln.pk0122.d3:5p5 5' end, mRNA sequence.

ACCESSION CA654430

VERSION CA654430

KEYWORDS EST.

SOURCE CA654430.1 GI:25232955

ORGANISM Triticum aestivum (bread wheat)

REFERENCE Triticum aestivum

AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Triticum.

Tingley,S.V., Powell,W., Wolters,P., Dolan,M., Hainey,C., Yuan,Z., Miao,G., Caraher,N. and Hanafey,M.K.

DuPont Wheat cDNA Sequence

Unpublished (2002)

Contact: Scott V. Tingley

Crop Genetics

E. I. DuPont de Nemours and Company

1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA

Tel: 302-631-2602

Fax: 302-631-2607

Email: Scott.V.Tingley@USA.dupont.com

PCR product - clone not available

Seq primer: M13.

Location/Qualifiers

1..1811

/organism="Triticum aestivum"

/mol\_type="mRNA"

/db\_xref="taxon:4565"

/clone="wreln.pk0122.d3:5p5"

/tissue\_type="root"

/clone\_lib="wreln"

/note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2: XhoI; Wheat (Triticum aestivum L.) root; normalized from wreln library"

ORIGIN

Query Match 24.3%; Score 555.8; DB 14; Length 1811;

Best Local Similarity 62.7%; Pred. No. 1.1e-93;

Matches	1000;	Conservative	0;	Mismatches	574;	Indels	21;	Gaps	8;
QY	353	GACGTGGTCTCGACTCGCTCATCCCTTACATCGACGACCCCAAGGACCGGACCGCCGTT	412						
Db	70	GACGTGGGCTGGGGCTGGTATGGGGTGGTGGAGGACCCCTGGGACCGGACGCCATC	129						
QY	413	TCCAGAGTGTGCGACGCTGGTACGAGCTCGACTCGCTCACCCGCAAGCAGCTCACCATC	472						
Db	130	TGCTCTGGTCTGGCGCACTGGTGAAGGTGAGCGCTGACGCCGCAAGCACTCACCGTC	189						
QY	473	GCGCTCTGTACACACCAACCCCGGCTCGCTCCGCGCGCTTCCCGCACTCGAGTCG	532						
Db	190	GCCATGGCTACTCTCACCAACCCCGGCTCTTCCGCGCTTCCCGCGCTCGCTCGAGTCG	249						
QY	533	CTCAGCTCAAGGGCAAG-CCCGAGCGCCATGTTCACTTGATACCCGA-GGATTGGG	590						
Db	250	CTCAGCTCAAGGGCAAGCCCGCGCTTCCATGTTCAACCTCATCCCGAGGAGTGGG	309						
QY	591	GCGGACAGTCACTCCCTGGGTCAAGAGATTCTCAGTACTTCGATTGCTTCAAGAGCC	650						
Db	310	GGAGTCCGCTCGCCCTGGATCCGCGAGCTCTCCGCTCTCTCCACTTCTCAAGATGC	369						
QY	651	TCCACTTCGCGCGATGATGTCAGAGATTCGATTCGATTCAGAAATCTCGCTCGTGACCGG	710						
Db	370	TGCACCTCGCGCGATTATTGCTCCGACCAACCTCCCGCTGCTCGTGCGCGCAAGG	429						
QY	711	GTACAGTGTTCACGCTCTCAAG-CTTGACAGTGTCTCGGTTTCAACCAACGATGGTCTT	769						
Db	430	GCCAAATGCTGTCTCCCTCGAGCCCTGGACCGCTGCTCGGTTTCTCCAACTCTTCCCTC	489						
QY	770	TTCATATCGGTCGCTTTTCAAGAGTTTAAAGAGTCTTCTTTTGGAGGAAAGCTCAATT	829						
Db	490	GGTCTCTCGCGGTGNTGGAGAACTGGAAACGTTCTTTCTTGAAGAAAGTCTGTT	549						
QY	830	CTTGAAGG---ACGAGAAATGGCTPACAGAGCTTGTCTTGAATAATACAGTCTTGAG	886						
Db	550	GCTGAGAAAGAAATGATGAATGGCTCCGTGAGCTTGCTTACCAGCAATACTGTCTCTGAG	609						
QY	887	ACTCTCAATTTTCTGTGACAGACATGCTGTGTGAAGATTGAGAGACCTTGACATTTTA	946						
Db	610	AGCTGAATTTCTTCTGACGGA---TCTCAGGGCAATCCCTCGACATCTTCTCTCCCT	666						
QY	947	GCTAAAAATGGCCCAACTTAGTGTCTGTGAACTTACTGACTGTGAAATACTGATCTT	1006						
Db	667	GTGCGAAATGGCGAAGGCTGAAACTCTCAAGATTAGCGACTGTTCATGCTGACCTG	726						
QY	1007	GTGAATCTTTAAGCATGCTCTGCGTGGAGAGTTTGTGGAGGACCT-----AC	1060						
Db	727	GTGACCTGTTCGTACAGCAAAACACTACAGACTTTGTCTGGTGGTTCCTTTGATGAT	786						
QY	1061	AACGAGAACCAAGATACCTGCTATATCATTTACCAGCAAGTATGTGATTTGGT	1120						
Db	787	CAAGATCAAGGTGGGAATTATGTAATACTATTTCCTCTTCTCAGTACAGCACTTGAT	846						
QY	1121	TTAACAATATTTGGAAGATAGTGTGCCATTTGTGTTTCATGTTTGCAGCCGCTACTAAA	1180						
Db	847	TTGCTCTACATGGGACAAATGAGATGACATATTATTTCCATATGTTGTCACACTCAAG	906						
QY	1181	AAATTGGATCTCTTATGCAATGCTAGACACGAGATCAATTTGATTTTAATCCAAAGG	1240						
Db	907	AAGTTGGACCTTCAGTTTCACTTCTTACCACAGAGGATCACTGTCAAATTAGTCCAGCGC	966						
QY	1241	TGTCCAATCTGAGAGTCTTGGACAGAAATGTAATTGGAGATAGAGGTTTAGAGTT	1300						
Db	967	TGCCAAATCTAGAAGTTTGGAGGTGAGGATGTGATAGAGATCGAGGTTTAGAGTT	1026						
QY	1301	CTTGGTCGTTGTTGAAGAGCTTAAAGAGGCTTAGGATTGAAAGGGCGGATGATGATCAA	1360						
Db	1027	ATTGCGGGACCTGCAGAAATTTACAGCACTCAGAGTCCAGAGAGAGAGATGATGACCAA	1086						
QY	1361	GGA---ATGAGAGATGAAGAGGTACTGTGTCCCATAGAGGGCTAATAGCTTGTACAG	1417						
Db	1087	GGAGGTCTTGAGGACGAAAGGGTAGAGTGCACACAGTAGGATTGATGGCTGTAGCTGAA	1146						

## RESULT 6

CF807278/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

CF807278 560 bp mRNA linear EST 27-OCT-2003  
 psHB022K06f USDA-IPAFS:Expression of Phytophthora sojae genes  
 during infection and propagation Phytophthora sojae cDNA clone  
 SHB022K06 5, mRNA sequence.

CF807278.1 GI:37995689

EST.

Phytophthora sojae

Phytophthora sojae

Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;

Phytophthora.

1 (bases 1 to 560)

Tyler, B.

Tyler, B. Not Published

Unpublished (2003)

Contact: Tyler B

Tyler lab

VBI

1880 Pratt Dr., Blacksburg, VA 24061, USA

Tel: 540-231-7318

Email: bmtylet@vt.edu

PCR Primers

FORWARD: BK reverse primer

BACKWARD: BK reverse primer

Plate: 022 row: K column: 06

Seq primer: BK reverse primer

High quality sequence stop: 560.

Location/Qualifiers

1..560

/organism="Phytophthora sojae"

/mol\_type="mRNA"

/db\_xref="taxon:67593"

/clone="SHB022K06"



1837 TCTTAGGTAAGTGGTGGCAAGGTTATGGTGTATCTCCATCTGACGTGATCTTTTGGT 1896  
 361 TCTTAGGTAAGTGGTGGCAAGGTTATGGTGTATCTCCATCTGACGTGATCTTTTGGC 420  
 1897 AATGCTCGACCTTTTGGACATTTAGTATCTTCTAGAAAGGTTGCTTACGAATAC 1956  
 421 AATGCTCGACCTTTTGGACATTTAGTATCTTCTAGAAAGGTTGCTTACGAATAC 480  
 1957 CAATCCAGATGAGCTAGTGTGTGAGCATCTCTCATATTTCTGATATTTATCTCT 2016  
 481 CAATTCAGATGAGCGGTAGTGTGTGAGCATCTCTCATATTTCTGATATTTATCTCT 540  
 2017 TCACGGGACAGATCAGATTTTCCAGATCTTT 2050  
 541 TGCAGGGACAGATCAGATTTTCCAGATCTTT 574

RESULT 8  
 BQ133928  
 LOCUS  
 DEFINITION  
 san55a01.y1 Gm-cl052 Glycine max cDNA clone SOYBEAN CLONE ID:  
 Gm-cl052-3458 5' similar to TR:004197 004197 HYPOTHETICAL PROTEIN.  
 ; mRNA sequence.  
 BQ133928  
 BQ133928.1 GI:20207839  
 EST.  
 SOURCE  
 Glycine max (soybean)  
 ORGANISM  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.  
 1 (bases 1 to 538)  
 Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V.,  
 Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,  
 Beck, C., Wylie, T., Underwood, K., Streptoe, M., Theising, B., Allen, M.,  
 Bowers, J., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,  
 Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,  
 McCann, R., Waterston, R. and Wilson, R.  
 Public Soybean EST Project  
 Unpublished (1999)  
 Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: estowatson.wustl.edu  
 This clone is available through: ResGen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntsville, AL 35801 For further information  
 call: (800)-533-4363 or contact: ccu@resgen.com web site:  
 www.resgen.com  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 421.  
 Location/Qualifiers  
 1..538  
 /organism="Glycine max"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3847"  
 /clone="SOYBEAN CLONE ID: Gm-cl052-3458"  
 /tissue\_type="whole seedlings of greenhouse grown plants"  
 /dev\_stage="1 week old"  
 /lab\_host="DH10B"  
 /clone\_lib="Gm-cl052"  
 /note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:  
 XhoI; The Harosoy NIL was constructed and seed was  
 provided by Dr. J. Specht, University of Nebraska  
 (Shoemaker and Specht, 1995). The cDNA library was  
 constructed from mRNA isolated from whole seedlings of 1  
 week old greenhouse grown plants. Complementary DNA was  
 synthesized from mRNA using a primer consisting of a  
 poly(dT) sequence with a XhoI restriction site and a 3'  
 anchor. EcoRI adapters were ligated to the blunt-ended

cDNA fragments followed by XhoI digestion. The cDNA  
 fragments were directionally cloned into the EcoRI-XhoI  
 restriction site of the pBluescript vector. The ligated  
 cDNA fragments were transformed into DH10B host cells  
 (GibcoBRL). The library was constructed in cooperation  
 with Dr. Paul Keim's laboratory at Northern Arizona  
 University."

ORIGIN

Query Match 23.5%; Score 538; DB 13; Length 538;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-90;  
 Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 307 ATCCAAATATGACGGAGGAACGGAAGTCCGGAAGACACGTGTGTGTCGAGCTGCTCTCGA 366  
 DB 1 ATCCAAATATGACGGAGGAACGGAAGTCCGGAAGACACGTGTGTGTCGAGCTGCTCTCGA 60  
 QY 367 CTGCGTCATCTCCCTTACATCTGACGACCCCAAGACCGCGCTTCCAGGTGTGTG 426  
 DB 61 CTGCGTCATCTCCCTTACATCTGACGACCCCAAGACCGCGCTTCCAGGTGTGTG 120  
 QY 427 AGCTGTGTACGAGTCGATCTGCTCACCGGACGACGTCACCATCGGCTCTGTACAC 486  
 DB 121 ACGCTGTGTACGAGTCGATCTGCTCACCGGACGACGTCACCATCGGCTCTGTACAC 180  
 QY 487 CACCAACCCCGGCTCGCTCCGCGCGGCTTCCCGCACCTCGAGTCGCTCAAGCTCAAGGG 546  
 DB 181 CACCAACCCCGGCTCGCTCCGCGCGGCTTCCCGCACCTCGAGTCGCTCAAGCTCAAGGG 240  
 QY 547 CAAAGCCCGGAGCGCGCAATGTTCAACTTGATATCCCGAGGATTTGGGGCGGACAGTCTCC 606  
 DB 241 CAAAGCCCGGAGCGCGCAATGTTCAACTTGATATCCCGAGGATTTGGGGCGGACAGTCTCC 300  
 QY 607 CTGGGTCAAGAGATTTCTCAGTACTTTCGATTCCTCAAGAGCTCCACTTCGCGCGCAT 666  
 DB 301 CTGGGTCAAGAGATTTCTCAGTACTTTCGATTCCTCAAGAGCTCCACTTCGCGCGCAT 360  
 QY 667 GATTGTCAAGGATTTCCGATCTTTCAGATCTTCCGCTGTGACCGCGGTTCAGTCTCTTCAACG 726  
 DB 361 GATTGTCAAGGATTTCCGATCTTTCAGATCTTCCGCTGTGACCGCGGTTCAGTCTCTTCAACG 420  
 QY 727 TCTCAAGCTTGACAGTGTCTCCGTTTACCAACCGATGTCTTTTCCATATCGGTGCTT 786  
 DB 421 TCTCAAGCTTGACAGTGTCTCCGTTTACCAACCGATGTCTTTTCCATATCGGTGCTT 480  
 QY 787 TTGCAAGATTTTAAAGATCTTCTTTTTCGAGGAAGTCAATTTCTTGAGAGACGCGA 844  
 DB 481 TTGCAAGATTTTAAAGATCTTCTTTTTCGAGGAAGTCAATTTCTTGAGAGACGCGA 538

RESULT 9  
 BI967384/c  
 LOCUS  
 DEFINITION  
 Gm830001B20C10 Gm-r1083 Glycine max cDNA clone Gm-r1083-140 3',  
 mRNA sequence.  
 BI967384  
 BI967384.1 GI:16341789  
 EST.  
 SOURCE  
 Glycine max (soybean)  
 ORGANISM  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.  
 1 (bases 1 to 561)  
 Vodkin, L., Keim, P., Shoemaker, R., Retzel, B., Khanna, A., Coryell, V.,  
 Erpelting, J., Raph, C., Shoop, E., Pardinas, J., Liu, L. and Lewin, H.  
 A Functional Genomics Program for Soybean (NSF 9872565)  
 Unpublished (1999)  
 Other ESTs: AW102154 corresponding to Gm-cl009-1054 (5')  
 Contact: Vodkin, L.O., PI, A Functional Genomics Program for  
 Soybean (NSF 9872565)  
 Lewin, H. A., Director, Keck Center for Comparative and Functional

Genomics University of Illinois Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA Tel: (217) 244-6147 Fax: (217) 333-4582 Email: l-vodkin@uiuc.edu This clone is available through: Incyte Genomics, 4633 World Parkway Circle St. Louis, Missouri 63134. Phone (800) 430-0030 or (314) 427-3222 FAX: (314) 427-3324. Web site: http://www.incyte.com/reagents/catalog.jsp?page=clones/collaboratio n/index Seq primer: 5'-TTTTTTTTTTTTTTTTT(A/C/G)-3'.		Genomics University of Illinois Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA Tel: (217) 244-6147 Fax: (217) 333-4582 Email: l-vodkin@uiuc.edu This clone is available through: Incyte Genomics, 4633 World Parkway Circle St. Louis, Missouri 63134. Phone (800) 430-0030 or (314) 427-3222 FAX: (314) 427-3324. Web site: http://www.incyte.com/reagents/catalog.jsp?page=clones/collaboratio n/index Seq primer: 5'-TTTTTTTTTTTTTTTTT(A/C/G)-3'.	
FEATURES		FEATURES	
Location/Qualifiers		Location/Qualifiers	
1..561 /organism="Glycine max" /mol_type="mRNA" /db_xref="taxon:3847" /clone="Gm-r1083-140" /clone_lib="Gm-r1083" /note="The library Gm-r1083 is a sequence-driven, reracked set of 4,992 clones selected from cDNA libraries from various tissues and stages of development of soybean. It represents 1117 sequences from the progenitor library Gm-cl009 (from mature roots of 2 month old greenhouse grown 'Williams' soybean plants); 820 sequences from the progenitor library Gm-cl013 (from 2 to 3 week old whole plants of Williams); and 3055 sequences from library Gm-cl028 (from 'Supernod' plants whose seedlings were inoculated with Bradyrhizobium japonicum, courtesy of Dr. Gary Stacey). The 5' ESTs of the source clones from the different progenitor libraries was used to select singletons, or a representative of each contig, which were reracked to form library Gm-r1083. The cDNA clones of the reracked Gm-r1083 library were then sequenced at the 3' end. The contig analysis to select unique genes was performed by the laboratory of Ernest Retzel, Center for Computational Genomics and Bioinformatics, University of Minnesota, http://web.ahc.umn.edu/biodata/nfsioy/. Reracking was performed by Incyte Genomics, St. Louis, http://www.incyte.com, and 3' sequencing by the Keck Center for Comparative and Functional Genomics, University of Illinois, http://www.lie.uiuc.edu/biotech/keck.html. Note: The corresponding 5' EST from each clone in the Gm-r1083 library is listed in the 'OTHER EST' field. The detailed information on the source library for each clone can also be obtained by referring to the Incyte Genomics clone ID of the original cDNA library that is also listed under 'OTHER EST'."		1..541 /organism="Phytophthora sojae" /mol_type="mRNA" /db_xref="taxon:67593" /clone="SHB022P18" /tissue_type="mycelium" /cell_line="P6497" /dev_stage="48 hr. post infection stage" /lab_host="Soybean plant" /clone_lib="USDA-IPAFS:Expression of Phytophthora sojae genes during infection and propagation" /note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"	
ORIGIN		ORIGIN	
Query Match 23.3%; Score 532.2; DB 12; Length 561; Best Local Similarity 98.6%; Pred. No. 3e-89; Matches 545; Conservative 0; Mismatches 7; Indels 1; Gaps 1;		Query Match 23.0%; Score 526; DB 14; Length 541; Best Local Similarity 100.0%; Pred. No. 4.3e-88; Matches 526; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1651	AGGCCTTGTTACATTCGACATACAGTCCAAATGTGAGATGATGCTCTGTTATGT	1710
DB	561	AGGCCTTGTTACATTCGACATACAGTCCAAATGTGAGATGATGCTCTGTTATGT	502
QY	1711	GGGGAGTCTGATCGAGGGCTTTTGAGTTCGTAGGGGTGCTAGTCTTCAGAACT	1770
DB	501	GGGGAGTCTGATCGAGGGCTTTTGAGTTCGTAGGGGTGCTAGTCTTCAGAACT	442
QY	1771	TGAATGAGAGGGGTGTTATTTTTCAGTGAACGTGACATTCGTGTGGCTGCAACACAAT	1830
DB	441	TGAATGAGAGGGGTGTTATTTTTCAGTGAACGTGACATTCGTGTGGCTGCAACACAAT	382
QY	1831	GACTTCTTTAGGTACTGTGGTGCAGAGTTATGTGTATCTCCATCTGGAAGTATCT	1890
DB	381	GACTTCTTTAGGTACTGTGGTGCAGAGTTATGTGTATCTCCATCTGGAAGTATCT	322
QY	1891	TTTGGTAATGGCTCGACCCCTTTTGGAAACATTTAGTTGATTCCTCTAGAAAGTGGCTAC	1950
DB	321	TTTGGTAATGGCTCGACCCCTTTTGGAAACATTTAGTTGATTCCTCTAGAAAGTGGCTAC	262

```
1823 ACACAAATGACTCTCTCTAGGTAAGTCTGCTGGTGCAGAGTTATGGTGTATCTCCATCTGGA 1882
Db      |||||||
481  ACACAAATGACTCTCTCTAGGTAAGTCTGCTGGTGCAGAGTTATGGTGTATCTCCATCTGGA 422
QY      |||||||
1883  CQTGATCTTTTGGTAATGCTCGACCTTTTGGACATTTGAGTTGATTCCTTCTAGAAAG 1942
Db      |||||||
421  CQTGATCTTTTGGTAATGCTCGACCTTTTGGACATTTGAGTTGATTCCTTCTAGAAAG 362
QY      |||||||
1943  GTGGCTACCAATACCAATACCAATGAGTGTAGTTGTTGAGCATCTCTCTCATATCTTT 2002
Db      |||||||
361  GTGGCTACCAATACCAATACCAATGAGTGTAGTTGTTGAGCATCTCTCTCATATCTTT 302
QY      |||||||
2003  GCATATATCTTCTGACGGCAGAGATCAGATTTTCCAGATCTGTTGCTTGGAC 2062
Db      |||||||
301  GCATATATCTTCTGACGGCAGAGATCAGATTTTCCAGATCTGTTGCTTGGAC 242
QY      |||||||
2063  ACTGCCACATGCTTGCATACCTAGAGCCAGAGCTGTGTATATATATACCAAGTTTCTTTTG 2122
Db      |||||||
241  ACTGCCACATGCTTGCATACCTAGAGCCAGAGCTGTGTATATATATACCAAGTTTCTTTTG 182
QY      |||||||
2123  TTTTCTCTCTCCCTTTCATATGCTGTTCTATGTTCTCTGCTCTATTTGTAAGTTCAATTT 2182
Db      |||||||
181  TTTTCTCTCTCCCTTTCATATGCTGTTCTATGTTCTCTGCTCTATTTGTAAGTTCAATTT 122
QY      |||||||
2183  AGCAATATAGTCTTGTAAATGAGCTGTTTTCATTTGAAATCTGAAAGCTTCCCTTA 2242
Db      |||||||
121  AGCAATATAGTCTTGTAAATGAGCTGTTTTCATTTGAAATCTGAAAGCTTCCCTTA 62
QY      |||||||
2243  ACCCTATTGGCTCCCTTAAATGAGCTGTTTTCATTTGAAATCTGAAATCTGAAATCTGAAAT 2288
Db      |||||||
61  ACCCTATTGGCTCCCTTAAATGAGCTGTTTTCATTTGAAATCTGAAATCTGAAATCTGAAAT 16
```

```
ORIGIN
Query Match      22.5%; Score 515.6; DB 12; Length 538;
Best Local Similarity 97.4%; Pred. No. 3.8e-86;
Matches 524; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY      790 CAAGAGTTTAAGAGTCTCTGTTTGGAGGAAAGCTCAATCTTTGAGAGGACGAGAAATG 849
Db      1 CAGGAGTTTAAGAGTCTCTGTTTGGAGGAAAGCTCAATCTTTGAGAGGACGAGAAATG 60
QY      850 GCTACACGAGTGTGCTTGAATAACATGCTTCTGAGACTCTCAATTTTACTTGACAGA 909
Db      61 GTTACACGAGTGTGCTTGAATAACATGCTTCTGAGACTCTCAATTTTACTTGACAGA 120
QY      910 CATTCGCTTGTGAGATTCAGACCTTGAACCTTTTAGCTTAAATAATGCCCCAACTTAGT 969
Db      121 TATTGCTGTTGTAAGATTCAGACCTTGAACCTTTTAGCTTAAATAATGCCCCAACTTAGT 180
QY      970 GTCGTGAAACTTACTGACTGTGAAATATCTGGAATCTTGTGAATCTTTAAAGCATGCTC 1029
Db      181 GTCGTGAAACTTACTGACTGTGAAATATCTGGAATCTTGTGAATCTTTAAAGCATGCTC 240
QY      1030 TGGCTGGAAGAGTGTGTTGGAGGACCTTACACGAGGAAACCAAGAAATGATCTGCTAT 1089
Db      241 TGGCTGGAAGAGTGTGTTGGAGGACCTTACACGAGGAAACCAAGAAATGATCTGCTAT 300
QY      1090 ATCAATACAGCAAGTATGTCGATGGTGTAAACATATATTGGAAGAAATGAGTTGCC 1149
Db      301 ATCAATACAGCAAGTATGTCGATGGTGTAAACATATATTGGAAGAAATGAGTTGCC 360
QY      1150 CATTTGTTTCATGTTTGCAGCCCTACTTAAATAATGGAATCTCTCTATGCAATGCTAGA 1209
Db      361 CATAGTTTCACTGTTTGCAGCCCTACTTAAATAATGGAATCTCTCTATGCAATGCTAGA 420
QY      1210 CACGAGGATCATGTTATGTTAATCCAAAGGTCCAAATCTGGAAGTCTTGGAGCAAG 1269
Db      421 CACGAGGATCATGTTATGTTAATCCAAAGGTCTCCAAATCTGGAAGTCTTGGAGCAAG 480
QY      1270 GAATGTAATTTGGAGATAGAGGTTTAGAGGTTCTTGGTCTGTTGTTGTAAGAGGCTTAAA 1327
Db      481 GAATGTAATTTGGAGATAGAGGTTTAGAGGTTCTTGGTCTGTTGTTGTAAGAGGCTTAAA 538
```

```
RESULT 12
LOCUS      CA820349
DEFINITION      550 bp mRNA linear EST 09-DEC-2002
saub87g05.y1 Gm-cl048 Glycine max cDNA clone SOYBEAN CLONE ID:
Gm-cl048-3009 5' similar to TR:004197 004197 HYPOTHETICAL PROTEIN.
i. mRNA sequence.
ACCESSION      CA820349
VERSION         CA820349.1
KEYWORDS        GI:26269286
SOURCE          EST.
               Glycine max (soybean)

RESULT 12
LOCUS      CA820349
DEFINITION      538 bp mRNA linear EST 30-NOV-2001
bai70d03.y1 Gm-cl068 Glycine max cDNA clone GENOME SYSTEMS CLONE
ID: Gm-cl068-3990 5' similar to TR:004197 004197 HYPOTHETICAL
PROTEIN. i. mRNA sequence.
ACCESSION      BI974579
VERSION         BI974579.1
KEYWORDS        GI:16348984
SOURCE          EST.
               Glycine max (soybean)
ORGANISM        Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 538)
Shoemaker,R., Keim,P., Vodkin,L., Rzepeljing,J., Coryell,V.,
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,R., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schuck,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCaun,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 410.
Location/Qualifiers
1..538
/organism="Glycine max"
/mol_type="mRNA"
```

QY	1832	ACTCTCTCTTAGGTACTTGTGGGTGCAAGGTATAGTGTGTATCTCCATCTGACGTGATCTT	1891
Db	241	ACTTCTCTTAGGTACTTGTGGGTGCAAGGTATAGTGTGTATCTCCATCTGACGTGATCTT	300
QY	1892	TTGGTAAATGGCTCGACCCCTTTTGGAAACATGTAGTTGATTCCTTCTAGAAAGGTGGCTACG	1951
Db	301	TTGGCAATGGCTCGCCCTTTTGGAAACATGTAGTTGATTCCTTCTAGAAAGGTGGCTATG	360
QY	1952	AATACCAATCCAGATCAGACTGTAGTTGTTGAGCATCTGCTCATATCTTTGCATATTAT	2011
Db	361	AATACCAATTCAGATGAGACGGTAGTTGTTGAGCATCTGCTCATATCTTTGCATATTAT	420
QY	2012	TCTCTTGCAGGGCAGAGATCAGATTTTCCAGATACATGTTGTGCTTTGGACACCTGCCACA	2071
Db	421	TCTCTTGCAGGGCAGAGATCAGATTTTCCAGATACATGTTGTGCTTTGGACACCTGCCACA	480
QY	2072	TGGCTTGATACCTAGAGGCGAGAGCTGTGTATATACAGTTTCTTTTGTGTTTCTTC	2131
Db	481	TGGCTTGACACCTTAGAGGCGAGAGCTGTGTATATACAGTTTCTTTTGTGTTTCTTC	527
QY	2132	TCCCTCTTCATATGCTGTTTCTA	2154
Db	528	TCCCTCTTCATATGTTGTTTCTA	550
RESULT 13			
AI855554			
LOCUS			
DEFINITION	AI855554.1 GI:5508996 493 bp mRNA linear EST 30-NOV-2001		
	8220801.Y1 Gm-cl013 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:		
	Gm-cl013-1225 5' similar to TR:004197 004197 HYPOTHETICAL PROTEIN.		
ACCESSION	AI855554		
VERSION	AI855554.1		
KEYWORDS	EST.		
SOURCE	Glycine max (soybean)		
ORGANISM	Glycine max		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
	rosids; eurosids I; Fabiales; Fabaceae; Papilionoideae; Phaseoleae;		
	Glycine.		
REFERENCE	1 (bases 1 to 493)		
AUTHORS	Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V.,		
	Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,		
	Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,		
	Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N.,		
	Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,		
	McCann, R., Waterston, R. and Wilson, R.		
	Public Soybean EST Project		
TITLE	Public Soybean EST Project		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Shoemaker R/Public Soybean EST Project		
	Public Soybean EST Project		
	Washington University School of Medicine		
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA		

Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available through: ResGen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntsville, AL 35801 For further information  
 call: (800)-533-4363 or contact via email: ccu@resgen.com  
 Insert Length: 1449 Std Error: 0.00  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 430.  
 Location/Qualifiers  
 1..493  
 /organism="Glycine max"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl013-1225"  
 /tissue\_type="whole seedlings, 2-3 week old seedlings,  
 greenhouse grown"  
 /lab\_host="XL10-Gold"  
 /clone\_lib="Gm-cl013"



/note="Vector: pBluescript II XR; Site 1: EcoRI; Site 2: XhoI; This cDNA library was constructed from mRNA isolated from whole seedlings of 2-3 week old greenhouse grown plants. The cDNA library was prepared using the Stratagene pBluescript II XR cDNA library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into XL10-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelding."

ORIGIN		Query Match	21.4%;	Score 489;	DB 9;	Length 493;
		Best Local Similarity	100.0%;	Pred. No. 3.5e-81;		
		Matches 489;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1007	GTGAATCTTTAAGCATGCTCTGCGCTGGAAGAGTTTGTGAGGCGACCTACACGAG	1066			
Db	5	GTGAATCTTTAAGCATGCTCTGCGCTGGAAGAGTTTGTGAGGCGACCTACACGAG	64			
QY	1067	GAACAGAAAGATCTGCTGTATATCATACAGCAAGTTATGTGATGGGTTTAAACA	1126			
Db	65	GAACAGAAAGATCTGCTGTATATCATACAGCAAGTTATGTGATGGGTTTAAACA	124			
QY	1127	TATATTGAAAGATGAGTTGCCATTTGTTTCATGTTTGCAGCGGTACTAAAAAATTG	1186			
Db	125	TATATTGAAAGATGAGTTGCCATTTGTTTCATGTTTGCAGCGGTACTAAAAAATTG	184			
QY	1187	GATCTCTTATGCAATGCTAGACAGGAGTATCTGTTTAAATCAAGGTTGTCCA	1246			
Db	185	GATCTCTTATGCAATGCTAGACAGGAGTATCTGTTTAAATCAAGGTTGTCCA	244			
QY	1247	AATCTGGAAGTCTTGAGCAAGGATGTAATTTGAGATAGAGGTTAGAGTTCTTGGT	1306			
Db	245	AATCTGGAAGTCTTGAGCAAGGATGTAATTTGAGATAGAGGTTAGAGTTCTTGGT	304			
QY	1307	CGTTGTTGTAAGAGGCTTAAAGAGGCTTAGGATTTGAAGGGCGGATGATCAAGGAATG	1366			
Db	305	CGTTGTTGTAAGAGGCTTAAAGAGGCTTAGGATTTGAAGGGCGGATGATCAAGGAATG	364			
QY	1367	GAGGATGAAGAGGTTACTGTGCTCCATAGAGGCTTAATAGCCTTGTACAGGCTGTTC	1426			
Db	365	GAGGATGAAGAGGTTACTGTGCTCCATAGAGGCTTAATAGCCTTGTACAGGCTGTTC	424			
QY	1427	GAGCTTGATATGCTGCTGTTATGTTGCTGATATTACAAATGCATCTCTGGAACATATT	1486			
Db	425	GAGCTTGATATGCTGCTGTTATGTTGCTGATATTACAAATGCATCTCTGGAACATATT	484			
QY	1487	GGAACCTCAC	1495			
Db	485	GGAACCTCAC	493			

RESULT 14  
BQ630213  
LOCUS  
DEFINITION BQ630213.1 555 bp mRNA linear EST 21-OCT-2002  
sap99d04.y1 Gm-c1045 Glycine max cDNA clone SOYBEAN CLONE ID:  
Gm-c1045-3248 5' similar to TR:004197 004197 HYPOTHETICAL PROTEIN.  
i, mRNA sequence.  
ACCESSION BQ630213  
VERSION BQ630213.1 GI:21677862  
KEYWORDS EST.  
SOURCE Glycine max (soybean)  
ORGANISM Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.  
REFERENCE 1 (bases 1 to 555)

AUTHORS

Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Corvelli,V.,  
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,  
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,  
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,  
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,  
McCann,R., Waterston,R. and Wilson,R.  
Public Soybean EST Project  
Unpublished (1999)

TITLE  
JOURNAL  
COMMENT

Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu  
This clone is available through: ResGen, Invitrogen Corp. 2130  
South Memorial Parkway Huntsville, AL 35801 For further information  
call: (800)-533-4363 or contact: ccu@resgen.com web site:  
www.resgen.com  
Seq primer: -40RP from Gibco  
High quality sequence stop: 434.  
Location/Qualifiers  
1. .555  
/organism="Glycine max"  
/mol\_type="mRNA"  
/db\_xref="taxon:3847"  
/clone="SOYBEAN CLONE ID: Gm-c1045-3248"  
/tissue\_type="Hypocotyl, 9-10 day old etiolated seedlings"  
/lab\_host="DH10B"  
/clone\_lib="Gm-c1045"  
/note="Vector: pBluescriptII SK+; Site 1: EcoRI; Site 2:  
XhoI; This cDNA library was constructed from mRNA isolated  
from etiolated hypocotyl tissue of 9-10 day old seedlings  
of the cultivar Williams 82. Complementary DNA was  
synthesized from mRNA using a primer consisting of a  
poly(dT) primer with a XhoI restriction site. EcoRI  
adapters were ligated to the blunt-ended cDNA fragments  
followed by digestion with EcoRI and XhoI. The cDNA  
fragments were directionally cloned into the EcoRI-XhoI  
restriction site of the pBluescript vector. The ligated  
cDNA fragments were transformed into DH10B host cells  
(Gibco, BRL). This library was constructed by Dr. Randy  
Shoemaker."

FEATURES  
source

ORIGIN

Query Match		20.8%;	Score 476.2;	DB 13;	Length 555;
		Best Local Similarity	93.3%;	Pred. No. 8.7e-79;	
		Matches 525;	Conservative 0;	Mismatches 23;	Indels 15; Gaps 2;
QY	1623	ATCTCAGGCGTGGCGGGTTGACTGATGTAGGCTTGGTTACATTGACAAATACAGTCCAA	1682		
Db	6	AGCTCAGGCGTGGCGGGTTGACCGATGTAGGCTTGGTTACATTGACAAATACAGTCCAA	65		
QY	1683	ATGTGAGATGAGTGTGCTTGGTTATGTGGGGAGTCTGATGACGGGCTTTTGGAGTTGG	1742		
Db	66	ATGTGAGATGAGTGTGCTTGGTTATGTGGGGAGTCTGATGACGGGCTTTTGGAGTTGG	125		
QY	1743	CTAAGGGGTGCTCCTAGTCTTCAGAACTTGAATGAGAGGGTGTATTATTTTCAGTGAAC	1802		
Db	126	CTAAGGGGTGCTCCTAGTCTTCAGAACTTGAATGAGAGGGTGTATTATTTTCAGTGAAC	185		
QY	1803	GTGCACTTCTGCTGGTGCACACATTCACCTTCTTCTTGGTACTTGTGGGTGCAAGGTT	1862		
Db	186	GTGCACTTCTGCTGGTGCACACATTCACCTTCTTCTTGGTACTTGTGGGTGCAAGGTT	245		
QY	1863	ATGGTGTATCTCCATCTGACGCTGATCTTTTGGTATGGCTCGACCTTTTGGAAATGG	1922		
Db	246	ATGGTGTATCTCCATCTGACGCTGATCTTTTGGCAATGGCTCGCCCTTTTGGAAATGG	305		
QY	1923	AGTTGATTCTCTTACAGAAAGGTGGGTACGAAATACCAATCCAGATGAGACTGTAGTTG	1982		
Db	306	AGTTAATCTCTTCTAGAAAGGTGGGTATGAATACCAATTCAGATGAGCGGTAGTTGTTG	365		





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 22:56:50 ; Search time 77 Seconds

(without alignments)  
2223.686 Million cell updates/sec

Title: US-10-009-791-22

Perfect score: 3180

Sequence: 1 TKTSAFLFTLSLRNMTTE.....QRSDFPVTVPPLDTATCVDT 606

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A Geneseq 29Jan04.\*

1: geneseqp1980s.\*

2: geneseqp1990s.\*

3: geneseqp2000s.\*

4: geneseqp2001s.\*

5: geneseqp2002s.\*

6: geneseqp2003as.\*

7: geneseqp2003bs.\*

8: geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3180	100.0	606	3 AAB23458	Aab23458 Soybean C
2	2126.5	66.9	592	3 AAB23466	Aab23466 Arabidops
3	1854.5	58.3	597	3 AAB23457	Aab23457 Rice COI1
4	1245.5	39.2	429	3 AAB23456	Aab23456 Corn COI1
5	818	25.7	585	3 AAG30234	Aag30234 Arabidops
6	812.5	25.6	236	3 AAB23449	Aab23449 Rice COI1
7	812.5	25.6	236	3 AAF01920	Aap01920 O. sativa
8	716.5	22.3	221	3 AAB23461	Aab23461 Wheat COI
9	678	21.3	522	3 AAG30235	Aag30235 Arabidops
10	613.5	19.3	177	3 AAF01922	Aap01922 T. aestiv
11	613.5	19.3	177	3 AAB23451	Aab23451 Wheat COI
12	579	18.2	490	3 AAG30236	Aag30236 Arabidops
13	555	17.5	108	3 AAB23450	Aab23450 Soybean C
14	555	17.5	108	3 AAF01921	Aap01921 G. max CO
15	476	15.0	342	3 AAG26270	Aag26270 Arabidops
16	475.5	15.0	194	3 AAB23448	Aab23448 Corn COI1
17	460	14.5	186	3 AAB23455	Aab23455 Corn COI1
18	460	14.5	186	3 AAF01926	Aap01926 Z. mays C
19	303.5	9.5	242	3 AAG26272	Aag26272 Arabidops
20	256.5	8.1	610	3 AAG12663	Aag12663 Arabidops
21	250.5	7.9	159	3 AAB23459	Aab23459 Wheat COI
22	249.5	7.8	610	3 AAG52202	Aag52202 Arabidops
23	239	7.5	243	3 AAG26271	Aag26271 Arabidops
24	213	6.7	405	3 AAG43539	Aag43539 Arabidops
25	202	6.4	721	4 AAG78364	Aag78364 GRR1 homo

## ALIGNMENTS

### RESULT 1

AAB23458  
ID AAB23458 standard; protein; 606 AA.

XX  
AC AAB23458;

XX  
DT 12-FEB-2001 (first entry)

XX  
DE Soybean COI1 protein #2.

XX  
KW Corn; rice; soybean; wheat; ATHCOR1; LLS1; disease resistance;  
KW A. thaliana coronatine-induced; COI1; herbicides; EST;  
KW expressed sequence tag; sg94c.pk003.k23.fis.

XX  
OS Glycine max.

XX  
PN WO200068406-A2.

XX  
PD 16-NOV-2000.

XX  
PF 03-MAY-2000; 2000WO-US011956.

XX  
PR 07-MAY-1999; 99US-0133041P.

XX  
PA (DUPO) DU PONT DE NEMOURS & CO E I.

XX  
PI Calmi PG, Famodu OO, Lee J, Miao G, Maxwell CA;

XX  
DR WPI; 2000-687649/67.

XX  
DR N-PSDB; AAA95063.

XX  
PT New nucleic acid sequences encoding new disease resistance factors,  
PT useful for producing plants with increased resistance to pathogens and to  
PT screen for herbicides.

XX  
PS Claim 10; Fig 1; 74pp; English.

XX  
CC Expressed sequence tags (ESTs) from rice, wheat, soybean and corn cDNA  
CC libraries were found to be similar to cDNA encoding the Arabidopsis  
CC thaliana COI1 protein and the Zea mays LLS1 protein. The present sequence  
CC is protein encoded by cDNA from the soybean clone sg94c.pk003.k23.fis  
CC which is homologous to COI1. The COI1 and LLS1 DNAs of the invention may be used  
CC to alter the expression of COI1 and LLS1 protein in cells, particularly  
CC to produce transgenic plants with increased systemic resistance to a wide  
CC range of pathogens. COI1 and LLS1 proteins may be used to identify  
CC inhibitors of these proteins, which may be useful as herbicides

Abb71219 Drosophil  
Aab23460 Wheat COI  
Aag78365 Arabidops  
ADJ10600 Human nov  
Aab92961 Human pro  
Aab92961 Human pro  
Aab92961 Human pro  
Abb64154 Drosophil  
Abo07181 Human p53  
Aab40624 Human ORF  
Aay83087 F-box pro  
Aab48305 Human ZF2  
Aao22473 Human F-2  
Aab48290 Human ZF1  
Aay02274 A F-box p  
Aao8046 Human ful  
Aag39654 Human F-b  
Aag43540 Arabidops  
Aag70904 Arabidops  
Aay83090 F-box pro

26 196 6.2 699 4 ABB71219  
27 192 6.0 134 3 AAB23460  
28 187.5 5.9 628 4 AAG78365  
29 179 5.6 422 5 ABJ10600  
30 179 5.6 436 6 ABR43242  
31 175 5.5 423 4 AAB92961  
32 173 5.4 423 4 AAB92961  
33 172.5 5.4 464 4 ABB64154  
34 172 5.4 423 6 ABO07181  
35 172 5.4 483 3 AAB40624  
36 172 5.4 483 3 AAY83087  
37 172 5.4 483 4 AAB48305  
38 172 5.4 483 5 AAO22473  
39 170.5 5.4 466 4 AAB48290  
40 164.5 5.2 456 2 AAY02274  
41 164.5 5.2 456 4 AAO8046  
42 164.5 5.2 456 7 AAG39654  
43 163.5 5.1 349 3 AAG43540  
44 161.5 5.1 693 4 AAG70904  
45 160.5 5.0 437 3 AAY83090

PD	16-NOV-2000.
XX	
XX	03-MAY-2000; 2000WO-US011956.
PF	
XX	
XX	07-MAY-1999; 99US-0133041P.
PR	
XX	(DUPO ) DU PONT DE NEMOURS & CO E I.
XX	
XX	Caimi PG, Famodu OO, Lee J, Miao G, Maxwell CA;
PI	
XX	WPI; 2000-687649/67.
XX	
XX	New nucleic acid sequences encoding new disease resistance factors,
PT	useful for producing plants with increased resistance to pathogens and to
PT	screen for herbicides.
PT	
XX	
XX	Example 3: Fig 1: 74pp: English.
PS	

XX	CC	Expressed sequence tags (ESTs) from rice, wheat, soybean and corn cDNA libraries were found to be similar to cDNA encoding the Arabidopsis thaliana COII protein and the Zea mays LLSI protein. The present sequence is the Arabidopsis thaliana COII protein. The COII and LLSI proteins are involved in disease resistance. The COII and LLSI DNAs of the invention may be used to alter the expression of COII and LLSI protein in cells, particularly to produce transgenic plants with increased systemic resistance to a wide range of pathogens. COII and LLSI proteins may be used to identify inhibitors of these proteins, which may be useful as herbicides
XX	CC	
XX	SQ	Sequence 592 AA;
		Query Match
		Best Local Similarity 66.9%; Score 2126.5; DB 3; Length 592;
		Matches 407; Conservative 72; Mismatches 97; Indels 9; Gaps 3
QY	19	ERNRTRV-----VDVLDDCVIPYIDDPKDRDAVSQCRRWYELDSLTKRHVTIALCY 73
DB	2	EDPIKECKLSCVATVDVIEQVMTYITDKDRDSASLCVRKFIDSETREHVTMALCY 61
QY	74	TTPPARLRPPPHUESLKLGKPRAMPNLIPDGWGHTVPWKVEISQFYDCCLKSLHPRR 133
DB	62	TATPDRLSRPNDRLSKLGKPRAMPNLIPENGGVYTPWVTETISNNLROLKSHVFR 121
QY	134	MIVXSDLIQNLRDGRHVHLHALKDCKSGFTTDLGLFHICRFCKSLRVLFLEESSILEKDG 193
DB	122	MIVSDLDRLAKARADDLETALKDKCSGFTTDGLLSIVTHCRKIKTILMESSSFSEKDG 181
QY	194	EWLHELALNNTVLETLPNYLTDAIVVKTEDLELLAKNCPNLVSVKLTDCEILLDVNPFKH 253
DB	182	KWLHELAQHTNSLEVINFYMTEFAKISPDKOLETIARNCSRSLSVKVGDFEILIELVGFKA 241
QY	254	ASALEEFCGGTYNEE--PERYSATSLPAKLCRLGLTVIGKNELTIVFMFAAVLKKLLDLL 310
DB	242	AANLEEFCCGSINEDIHGPEKYMWLVFPKLCRLGLSYGNPNEMPILPFFAAQIRKLDDL 301
QY	311	YAMLTDEHCLMIQRCPNLEVLRTNRVIGDRGLEVLGRCCCKLKRIRTERGDODQGWDE 370
DB	302	YALLETDHCTLIQKCPNLEVLRTNRVIGDRGLEVLQAQYCKQLKRLIRERGADEQQWDE 361
QY	371	BGTVSHRGILIASOCSLEYMAVYVSDITNASLEHIGTHLNKLCDFRVLVLDHEEKITD 430
DB	362	EGLVSQRGILIAAQCSLEYMAVYVSDITNESLESIGTYLNKLCDFRVLVLDREBRITD 421
QY	431	LPLDNGVRALLRGCDKLERFALYLRRGGITDVGLGYIGQYSNVWRMLLYGVGESDAGLL 490
DB	422	LFLDNGVRSLLIGCKLRRFAFLYRQGGLTDLGSLYIGQYSNVWRMLLYGVGESDEGLM 481
QY	491	EPFAKGCPSLOKLEMRCGLPFPSERALAVATOLTSLRYLMVQGVGVSPSGRDLIVMARPF 550
DB	482	EFSGPCNPLOKLEMRCGCC-FSERATAAATVKULPSRLYLMVQGVYRASMTQGDLMQMARP 540
QY	551	NTELIPSRKVATNTPDETWWVHEPHAILIAYTSLAGQRSDFPDPTV 595



Sequence 429 AA;

```
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146389P.
PR 02-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149923P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.

PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 25.7%; Score 818; DB 3; Length 585;
Best Local Similarity 34.6%; Pred. No. 3.4e-77; Indels 32; Gaps 12;
Matches 200; Conservative 109; Mismatches 237;

Qy 32 VLDVPIYDDPKORDAVSQVCRWYELDSLTRKHVTIALCYTTTPARLRFRPHLESK 91
Db 9 VLEHILSFIDSNEDNSVLSVCKSWPETERKTRKRVFVGNVCYAVPAVTRFRFEMRSLT 68
Qy 92 LKGPRAAMFNLIPEDEGWGHVTPWVKEISQVDFDCKLSLHFRMTVKSDLNLA---RDR 148
Db 69 LKGPHEADYNLVPDGGWGYAWPIEAMAAKSSLEIRMKRMVVTDECLEKIAASFKD- 127
Qy 149 GHVLHALKDKSCGFTDGLFHICRCKSLRVLFLERSSILEKDGEMHELALNNTVLET 208
Db 128 ---FKVLVLTSCBGFSTIDGIAIAATCRNLKVLRECEIVEDLGDWLSYFPESSTLSV 184
Qy 209 LNFYLTDIADVVKIRDLELLAKNCNPNVSVKLTDCBILD-LVNFFKHASALBEFCGTYNE 267
Db 185 LDFSCLD-SEVKISDLERLNSRSPNLKSLKLNPAVTLDDGLVSLRCAPOLTELGTGSPAA 243
Qy 268 E--PERYSAISLPAKLCR-----LGLTYIGKNELPIVFWFAVLKLLDLYAMLDTEHC 320
Db 244 QLKPEAFSKLSEAFSNCKQLQSLGLMDVLPYLPALYSVCPGLTSLNLSYATVRMPDLV 303
Qy 321 MLIORCPNLEVTNRNVIQDRGLVGRCKRKLRLRIERGDDQDQDEEGTVSHRGLI 380
Db 304 ELLRRCCKLQKLVMDLIEDKLEAVASYCKELRELVRFPSEPD--LDATNIPLEQGLV 361
Qy 381 ALSQGCSELEYMAVVSVDITNASLEHIGTHLKNLCDFRLVLD--HEKIDPLDNGVR 438
Db 362 FVSKGCRKLSVLYFCVQFTNAALFTIARKRPNLKCFELCVIEFPFAPDYKINEPLDKGFK 421
Qy 439 ALLRGCDKLRFPALYLRGGGLTDVGLGYIGQYSPNRWMLLYGVGESDAGLEFAKGCP 498
Db 422 AIAEGCRDLRLSV---SGLLSDKAFKYIGHAKKVRMLSIAGDSMLMLHLLSGCES 478
Qy 499 LQKLEMRGCLFFSERALAVAATQTLTSLRYLVQGYGVSPSGRDLVLMARPPFNWIELPSR 558
Db 479 LKLEIRDCP-FGDTALLLHAUKLETMSLWSSCFVSFGACKLLSOKMPLRNLNVEVI--- 534
Qy 559 KVATNTPDETUVVEHPA-HILAYISLAGQSDPDPTV 595
Db 535 ----DEHPPESRPSPVRIYIYRTVAGPRMDTPEFV 568

RESULT 6
AAB23449
ID AAB23449 standard; protein; 236 AA.
XX
AC AAB23449;
```

XX DT 12-FEB-2001 (first entry)  
XX DE Rice CO11 protein.  
XX KW Corn; rice; soybean; wheat; ATHCOR1; LLS1; disease resistance;  
XX DE A. thaliana coronatine-induced; CO11; herbicides; EST;  
XX KW expressed sequence tag; contig.  
XX OS Oryza sativa.  
XX FH Key Location/Qualifiers  
FT Misc-difference 115 /label= Unknown  
FT Misc-difference 219 /note= "Encoded by GG"  
XX PN W0200068406-A2.  
XX PD 16-NOV-2000.  
XX PF 03-MAY-2000; 2000WO-US011956.  
XX PR 07-MAY-1999; 99US-0133041P.  
XX PA (DUPO ) DU PONT DE NEMOURS & CO E I.  
XX PI Caimi PG, Famodu OO, Lee J, Miao G, Maxwell CA;  
XX DR WPI; 2000-687649/67.  
XX DR N-PSDB; AAA95054.  
XX PT New nucleic acid sequences encoding new disease resistance factors,  
XX PT useful for producing plants with increased resistance to pathogens and to  
XX PT screen for herbicides.  
XX PS Claim 10; Page 44-45; 74pp; English.  
XX CC Expressed sequence tags (ESTs) from rice, wheat, soybean and corn cDNA  
XX CC libraries were found to be similar to cDNA encoding the Arabidopsis  
XX CC thaliana CO11 protein and the Zea mays LLS1 protein. The present sequence  
XX CC is protein encoded by cDNA from the rice contig which is homologous to  
XX CC CO11. The CO11 and LLS1 proteins are involved in disease resistance. The  
XX CC CO11 and LLS1 DNAs of the invention may be used to alter the expression  
XX CC of CO11 and LLS1 protein in cells, particularly to produce transgenic  
XX CC plants with increased systemic resistance to a wide range of pathogens.  
XX CC CO11 and LLS1 proteins may be used to identify inhibitors of these  
XX CC proteins, which may be useful as herbicides  
XX SQ Sequence 236 AA;  
Query Match 25.6%; Score 812.5; DB 3; Length 236;  
Best Local Similarity 66.8%; Pred. No. 3e-77;  
Matches 157; Conservative 30; Mismatches 47; Indels 1; Gaps 1;  
QY 328 NLEVLTRNVIGDRGLEVLGRCKKRLRLRIGRDDQGDGMEDEGTVSHRGLIALSOGCS 387  
Db 1 NLLVLAVRNVIIGDRGLGVVADTKKQLRLRVERGDDDPGLQEEQGVSVGLTTVAVGCR 60  
QY 388 ELEVMAYVYSDITNASLEHIGTHLKNLCDPRLVLLDHEEKITDPLDNGVRALLRGCDKL 447  
Db 61 ELEVIAYVSDITNGALESIGTFCKNLCDPRLVLLDREERITDPLDNGVRALLXGCTKL 120  
QY 448 RRFALYLRGGGLTDVGLGYIGQYSPNVNRMMLGYGVEDSAGLLEFAKGCPSLOKLEMRGC 507  
Db 121 RRFALYLRPGGLSDTGLGYIGQYSGIIQYMLLNGVGETDGLTRFALGCENLRKLEIRSC 180  
QY 508 LFSERALAVAAQTLSRLVWQGVSPSGRDLVWARPFWNIIPSRKVAT 562  
Db 181 C-FSEQALARAIRSMPSRLRYVWQGVKASKTGHDLMLMARPFNIEPTPRRLVT 234  
RESULT 7

AAP01920  
ID AAP01920 standard; protein; 236 AA.  
XX AC AAP01920;  
XX DT 30-OCT-2001 (first entry)  
XX DE O. sativa CO11.  
XX KW CO11; rice; herbicide; disease-resistance factor.  
XX OS Oryza sativa.  
XX FH Key Location/Qualifiers  
FT Misc-difference 115 /label= "unknown  
FT Misc-difference 219 /note= "encoded by ANG"  
FT Misc-difference 219 /note= "encoded by GG"  
XX PN W0200068406-A2.  
XX PD 16-NOV-2000.  
XX PF 03-MAY-2000; 2000WO-US011956.  
XX PR 07-MAY-1999; 99US-0133041P.  
XX PA (DUPO ) DU PONT DE NEMOURS & CO E I.  
XX PI Caimi PG, Famodu OO, Lee J, Miao G, Maxwell CA;  
XX DR WPI; 2000-687649/67.  
XX DR N-PSDB; AAN02372.  
XX PT New nucleic acid sequences encoding new disease resistance factors,  
XX PT useful for producing plants with increased resistance to pathogens and to  
XX PT screen for herbicides.  
XX PS Claim 10; Page 44-45; 74pp; English.  
XX CC An isolated polynucleotide (I) comprising a sequence (Ia) encoding a  
XX CC polypeptide (II) of at least 60 amino acids (aa) from any of 18 specified  
XX CC sequences, given in the specification, or the complements of (Ia), is  
XX CC new. Herbicide. (I) encodes disease-resistance factor polypeptides (II)  
XX CC and can be used, in sense or antisense orientation, to alter the  
XX CC expression of (II) in cells, particularly to produce transgenic plants  
XX CC with increased systemic resistance to a wide range of pathogens. (I) or  
XX CC their fragments are also used for gene mapping (e.g. for plant breeding)  
XX CC and to identify loss of function mutants. Polypeptides (II) encoded by  
XX CC (I) are used to raise specific antibodies (for detection of (II)) and to  
XX CC design and/or identify specific inhibitors of (II), potentially useful as  
XX CC herbicides. This sequence is the rice CO11 protein created from a contig  
XX CC of xlr2.pk0027.b4, rlon.pk039.p14 and rlon.pk0047.c5 as described in the  
XX CC method of the invention  
XX SQ Sequence 236 AA;  
Query Match 25.6%; Score 812.5; DB 3; Length 236;  
Best Local Similarity 66.8%; Pred. No. 3e-77;  
Matches 157; Conservative 30; Mismatches 47; Indels 1; Gaps 1;  
QY 328 NLEVLTRNVIGDRGLEVLGRCKKRLRLRIGRDDQGDGMEDEGTVSHRGLIALSOGCS 387  
Db 1 NLLVLAVRNVIIGDRGLGVVADTKKQLRLRVERGDDDPGLQEEQGVSVGLTTVAVGCR 60  
QY 388 ELEVMAYVYSDITNASLEHIGTHLKNLCDPRLVLLDHEEKITDPLDNGVRALLRGCDKL 447  
Db 61 ELEVIAYVSDITNGALESIGTFCKNLCDPRLVLLDREERITDPLDNGVRALLXGCTKL 120  
QY 448 RRFALYLRGGGLTDVGLGYIGQYSPNVNRMMLGYGVEDSAGLLEFAKGCPSLOKLEMRGC 507  
Db 121 RRFALYLRPGGLSDTGLGYIGQYSGIIQYMLLNGVGETDGLTRFALGCENLRKLEIRSC 180

Qy	508	LFFSERALAVAAATOLTSRLYLWVGYSVPSGRDLLVMARFFWNIILPSKVAT	562
Db	181	C-FSEQALARAIRSMPSRLRYVVGYSKASKTGHDLMRLMARFFWNIILPRLVLT	234
RESULT 8			
AAB23461			
ID	AAB23461	standard; protein; 221 AA.	
XX	AC	AAB23461;	
XX	DT	12-FEB-2001 (first entry)	
XX	DE	Wheat CO11 protein #4.	
XX	KW	Corn; rice; soybean; wheat; ATHCOR1; LLS1; disease resistance;	
XX	KW	A. thaliana coronatine-induced; CO11; herbicides; EST;	
XX	KW	expressed sequence tag; wrln.pk0122.d3.fis.	
XX	OS	Triticum aestivum.	
XX	PN	WO200068406-A2.	
XX	PD	16-NOV-2000.	
XX	XX	03-MAY-2000; 2000WO-US011956.	
XX	XX	07-MAY-1999; 99US-0133041P.	
PA	(DUPO )	DU PONT DE NEMOURS & CO E I.	
PI	Caimi PG,	Fanodu OO, Lee J, Miao G, Maxwell CA;	
XX	WPI;	2000-687649/67.	
DR	N-PSDB;	AAA95066.	
XX	PT	New nucleic acid sequences encoding new disease resistance factors,	
PT	useful	for producing plants with increased resistance to pathogens and to	
PT	screen	for herbicides.	
XX	Claim	10; Page 66-67; 74pp; English.	
XX	Expressed	sequence tags (ESTs) from rice, wheat, soybean and corn cDNA	
CC	libraries	were found to be similar to cDNA encoding the Arabidopsis	
CC	thaliana	CO11 protein and the Zea mays LLS1 protein. The present sequence	
CC	is	protein encoded by cDNA from the wheat clone wrln.pk0122.d3.fis which	
CC	is	homologous to CO11. The CO11 and LLS1 proteins are involved in disease	
CC	resistance.	The CO11 and LLS1 DNAs of the invention may be used to alter	
CC	the	expression of CO11 and LLS1 protein in cells, particularly to produce	
CC	transgenic	plants with increased systemic resistance to a wide range of	
CC	pathogens.	CO11 and LLS1 proteins may be used to identify inhibitors of	
CC	these	proteins, which may be useful as herbicides	
XX	Sequence	221 AA;	
SQ	Query	Match	
	Best	Local Similarity 66.1%; Pred. No. 4.3e-67;	
	Matches	146; Conservative 22; Mismatches 44; Indels 9; Gaps 3;	
Qy	378	GLIALSQGSELEYMAVYSDITNASLEHIGTHKNLCDFRVLDDHEEKITDLPDNGV	437
Db	4	GLMAVAGGPDLEYMAVHVSDDITNALEAIGARKNLNDPRLVLLDREHITELPDNGV	63
Qy	438	RALLRGCDKLRRFALYLRGGLTDVGLGYIGQYSPNVRWMLLYGVGESDAGLLEPAKGP	497
Db	64	RALLRGCTKLRRFAFYVRPCALSDIGLSYVGFEFTKTVYMLNAGGSDGLLAFARGCP	123
Qy	498	SLQKLEMRGCLFFSERALAAVAATOLTSRLYLWVGYSVPSGRDLLVMARFFWNIILPS	557
Db	124	SLQKLELRSCC-FSERALAAVALQLKSLRYLWVGYSKASPTGDLMAVVRPFWNIIFIA-	181
Qy	558	RKATNTNPDETVEVHPAILAYLSLAGORSDFPDVTVPVL	598



PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	24-JUN-1999;	99US-0140991P.
PR	29-JUN-1999;	99US-0141287P.
PR	30-JUN-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	01-JUL-1999;	99US-0142055P.
PR	02-JUL-1999;	99US-0142390P.
PR	06-JUL-1999;	99US-0142803P.
PR	08-JUL-1999;	99US-0142920P.
PR	09-JUL-1999;	99US-0142977P.
PR	12-JUL-1999;	99US-0143542P.
PR	13-JUL-1999;	99US-0143624P.
PR	14-JUL-1999;	99US-0144005P.
PR	15-JUL-1999;	99US-0144086P.
PR	16-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	19-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145224P.
PR	23-JUL-1999;	99US-0145276P.
PR	26-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147260P.
PR	05-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0152363P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157533P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
Query Match 21.3%; Score 678; DB 3; Length 522;		
Best Local Similarity 33.7%; Pred. No. 2.2e-62;		
Matches 176; Conservative 96; Mismatches 219; Indels 32; Gaps 12;		
Qy	87	LESLKLGKPRAAFMNLPEDWGHVTPWVKELISQYFDCIKSLHFRMIVKDSLOJLA- 145
Db	1	MRSLLKKGKPHFADYNLPDGGWGYAWPWTEAMAKSSSLEETRMKRMVVTDECLEKIAA 60
Qy	146	--RDRGHVLHKLKDKCSGTTDGLFHIGRCKSLRVLFLEESSILEKDGELHLEALNN 203
Db	61	SFKD----FKVLVLTSCGFGFTDGIATAAATCRNLRVLELRCEIVEDLGGDWLSYFFESS 116

QY	204	TVLETLNFIYLDIAVVKIEDLELLAKNCPNLVSUKLTDCEILD-LVNFFKHASALEEFCG	262		
Db	117	TSLVSLDFSCILD-SEVKISDLERLVSRPNLSKLNPAVTLTDLGLVSLRCAPQLTELTG	175		
QY	263	GTYNEE--PERYSALSIPAKLCR-----LGLTYIGKNELPVVFMAAVLKLDLLIYAML	315		
Db	176	GSPAQLKPEAFSKLSEAFSNCKQLQSLSGLMDVLPVYLPAIYSCVPGLTSLNLSYATVR	235		
QY	316	TEDHMLIQRCPNLEVLNTRVIGDRGLEVLGRCKRLRLRIERGDDDDQGWEDDEGIVS	375		
Db	236	MPDLVELLRCSKQLKWMVMDLIEDKLEAVASYCKELRELRFVPEPD--LDATNIPLT	293		
QY	376	HRGLIALSQSCSELEYMAVYVSDITNASLEHIGTHLKNLCDFRLVLLD--HEEKITDPL	433		
Db	294	EQGLVFSVSGCRKLESVLYFCQFTNAALFTIARKPNLKCPRCLCVIEFFAPDYKTNEL	353		
QY	434	DNGVALLRGCDKLRREPALYLRGGITDVLGYIGQYSPNVRWMLLGYVGESDAGLLBFA	493		
Db	354	DGFKAIAGGCDLRLSV--SGLLSDKAFYIGKHAKVEMLSIAFAGSDMLHLL	410		
QY	494	KGCPSLQKLEMRGCLFFSERALAVAATQTSRYLWVQYGVSPSGRDLVWARPFWNIE	553		
Db	411	SGCESLKKLEIRDCP-FGDTALLEHAAKLETMRSLWMSGSCFVSFGACKLLSQMPRLNVE	469		
QY	554	LIPSRKVATNINDETVVVEHPA-HILAYYSLAGORSDFPDV	595		
Db	470	VI-----DEHPPESRPSESSVERIYIYRTVAGPRMDTPEFV	505		
RESULT 10					
ID	AAP01922	standard; protein; 177 AA.			
XX	AC	AAP01922;			
DT	30-OCT-2001	(first entry)			
XX	DE	T. aestivum COIL.			
XX	KW	COIL; wheat; herbicide; disease-resistance factor.			
XX	OS	Triticum aestivum.			
XX	Key	Location/Qualifiers			
FT	Misc-difference 84	/note= "encoded by GATA"			
XX	PN	WO200068406-A2.			
XX	PD	16-NOV-2000.			
XX	PF	03-MAY-2000; 2000WO-US011956.			
XX	PR	07-MAY-1999; 99US-0133041P.			
XX	PA	(DUPO ) DU PONT DE NEMOURS & CO E I.			
XX	PI	Caimi PG, Ramodu OO, Lee J, Miao G, Maxwell CA;			
XX	DR	WPI; 2000-687649/67.			
XX	DR	N-ESDB; AAN02374.			
XX	XX	New nucleic acid sequences encoding new disease resistance factors, useful for producing plants with increased resistance to pathogens and to screen for herbicides.			
XX	PS	Claim 10; Page 48-49; 74pp; English.			
XX	CC	An isolated polynucleotide (I) comprising a sequence (Ia) encoding a polypeptide (II) of at least 60 amino acids (aa) from any of 18 specified sequences, given in the specification, or the complements of (Ia), is new. Herbicide. (I) encodes disease-resistance factor polypeptides (II) and can be used, in sense or antisense orientation, to alter the			
CC	CC	expression of (II) in cells, particularly to produce transgenic plants with increased systemic resistance to a wide range of pathogens. (I) or their fragments are also used for gene mapping (e.g. for plant breeding) and to identify loss of function mutants. Polypeptides (II) encoded by (I) are used to raise specific antibodies (for detection of (II)) and to design and/or identify specific inhibitors of (II), potentially useful as herbicides. This sequence is the wheat COIL clone created from contigs wreln.pk0122.d3 and wlin.pk0018.f8 as described in the method of the invention			
XX	XX				
SQ	Sequence 177 AA;				
Query Match 19.3%; Score 613.5; DB 3; Length 177;					
Best Local Similarity 68.5%; Pred. No. 2.7e-56;					
Matches 122; Conservative 19; Mismatches 36; Indels 1; Gaps 1;					
QY	378	GLIALSQSCSELEYMAVYVSDITNASLEHIGTHLKNLCDFRLVLLDHEEKITDPLDNGV	437		
Db	1	GLMAVAGCPDLEYMAVHVSDITNALEAIGAFSKNLDNDFRLVLLDREVHITELDNGV	60		
QY	438	RALLRGCDKLRREPALYLRGGITDVLGYIGQYSPNVRWMLLGYVGESDAGLLBFAKGP	497		
Db	61	RALLRGCTKLRFPFVVRPGALSDLAFLXGFEFSTVYVYMLLGNAGSGDDGLLAFARXCP	120		
QY	498	SLQKLEMRGCLFFSERALAVAATQTSRYLWVQYGVSPSGRDLVWARPFWNIELI	555		
Db	121	SLQKLELRSCC-FSERALAAALQLKSURLYVWQYKASPTGTDLMAVVRPFWNIEFI	177		
RESULT 11					
ID	AAB23451	standard; protein; 177 AA.			
XX	AC	AAB23451;			
DT	12-FEB-2001	(first entry)			
XX	DE	Wheat COIL protein.			
XX	KW	Corn; rice; soybean; wheat; ATHCOR1; LLS1; disease resistance; A. thaliana coronatine-induced; COIL; herbicides; EST;			
XX	OS	Triticum aestivum.			
XX	Key	Location/Qualifiers			
FT	Misc-difference 85	/note= "Encoded by ATTG"			
FT	Misc-difference 89	/label= Unknown			
FT	Misc-difference 118	/label= Unknown			
XX	PN	WO200068406-A2.			
XX	PD	16-NOV-2000.			
XX	PF	03-MAY-2000; 2000WO-US011956.			
XX	PR	07-MAY-1999; 99US-0133041P.			
XX	PA	(DUPO ) DU PONT DE NEMOURS & CO E I.			
XX	PI	Caimi PG, Famodu OO, Lee J, Miao G, Maxwell CA;			
XX	DR	WPI; 2000-687649/67.			
XX	DR	N-PSDB; AAA95056.			
XX	XX	New nucleic acid sequences encoding new disease resistance factors, useful for producing plants with increased resistance to pathogens and to screen for herbicides.			
XX	PS	Claim 10; Page 47-48; 74pp; English.			

XX Expressed sequence tags (ESTs) from rice, wheat, soybean and corn cDNA  
CC libraries were found to be similar to cDNA encoding the Arabidopsis  
CC thaliana COI1 protein and the Zea mays L1S1 protein. The present sequence  
CC is protein encoded by cDNA from the wheat contig which is homologous to  
CC COI1. The COI1 and L1S1 proteins are involved in disease resistance. The  
CC COI1 and L1S1 DNAs of the invention may be used to alter the expression  
CC of COI1 and L1S1 protein in cells, particularly to produce transgenic  
CC plants with increased systemic resistance to a wide range of pathogens.  
CC COI1 and L1S1 proteins may be used to identify inhibitors of these  
CC proteins, which may be useful as herbicides  
XX  
SQ Sequence 177 AA;

Query Match 19.3%; Score 613.5; DB 3; Length 177;  
Best Local Similarity 68.5%; Pred No. 2.7e-56;  
Matches 122; Conservative 19; Mismatches 36; Indels 1; Gaps 1;

QY 378 GLIALSQGSELEYMAVYVSDITNASLEHIGTHLKNLDCFRVLVLDHEEKITDPLDNGV 437  
DB 1 GLMAVAGCPDLEYNAVHVSDITNALEAIGAFSKNLNDFRLVLLDREVHITELPLDNGV 60  
QY 438 RALLGCKLRFPALYLRGGTLDVGLGVIQYSPNRMILGYGESDAGLLEFAKQCP 497  
DB 61 RALLRGCTKLRFAFYVRPGALSDLAFLXIGBFSKTVRYMLGNAGSGDDGLAFARXCP 120  
QY 498 SLQKLEMRGCLFESERALAAVATQTLTSLRYLVQGYGSPSGRDLVLMARPPWNIELI 555  
DB 121 SLQKLEMRGCLFESERALAAVATQTLTSLRYLVQGYGSPSGRDLVLMARPPWNIELI 177

RESULT 12  
AAG30236  
ID AAG30236 standard; protein; 490 AA.  
XX  
AC AAG30236;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 36112.  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.

PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.

[illegible]



PR 25-MAR-1999; 99US-0136264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 17-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140931P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142820P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143342P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 19-JUL-1999; 99US-0144336P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144844P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145132P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.

Tue Apr 20 10:32:13 2004

```
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 15.0%; Score 476; DB 3; Length 342;
Best Local Similarity 36.6%; Pred. No. 3 3e-41;
Matches 113; Conservative 51; Mismatches 135; Indels 10; Gaps 5;

QY 32 VLDVPIYIDDPKDRDAVSQVRRWYELDSLTKKHVTIALCYTTTPARLRRRPHLESXK 91
Db 13 VLEHVFSTQLDKRNSVSLVCKSWYERWCKRKVFIGNCYAVSPATVIRRPKVRVSE 72

QY 92 LKGKPRAMFNLIPEWGGHVTWPVKEISQYDFCLSLHFRMIVKXDSLQNLARDRHV 151
Db 73 LKGKPHFADFNLVDPDGGVYVPIEAMSSSYTWLEELRKRMVVTDCLLEIAKSFKN- 131

QY 152 LHALKDKCGFTTDLGHIGRCKSLRVLFLEESSILEKDGWHLHAIANNVTLETNF 211
Db 132 FKVLVLSCEGFTDGLAIAATCRNLKELDLRESDDVSGHWSHFPDPTYSLSVSLNI 191

QY 212 YLTDAIAVKIEDLELLAKNCPNLVSKLTDCEILD-LVNFPHASALEBPCGGTYNEE-- 268
Db 192 SCL-ASEVSPSALERLVTRCPNLKSLNRAVPLEKIALTLQAPQLEELGTGGYTAVR 250

QY 269 PERYSALSIPAKLCR-----LGTIYICKNELPIVEMFAVLKKLDLYAMLDTEDHCLMI 323
Db 251 PDYISGLSVALSGCKELRCLSGFWDAVPAYLPVIVSCLRLLTINLSYATVQSYDLVKLL 310

QY 324 QRCNLEVL 332
Db 311 CQCPKLQRL 319
```

Search completed: April 20, 2004, 01:36:33  
Job time : 80 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: April 20, 2004, 01:34:16 ; Search time 25 seconds  
(without alignments)  
1251.414 Million cell updates/sec

Title: US-10-009-791-22  
Perfect score: 3180  
Sequence: 1 TKTSAPFLTSLRSNTTEE.....QRSDPDTVPVLDATCVDTI 606

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	164.5	5.2	456	3	US-09-172-841-51
2	164.5	5.2	456	4	US-08-951-621-51
3	127.5	4.0	1151	4	US-09-177-165A-31
4	119	3.7	483	4	US-09-904-615-154
5	113.5	3.6	380	4	US-09-172-841-53
6	113.5	3.6	380	4	US-08-951-621-53
7	112.5	3.5	435	2	US-08-531-439B-4
8	111.5	3.5	1874	4	US-09-331-403-2
9	110.5	3.5	579	4	US-09-325-932A-185
10	109	3.4	1495	3	US-08-462-467B-12
11	108	3.4	870	4	US-09-252-991A-28407
12	104	3.3	1074	4	US-09-004-838-111
13	103.5	3.3	980	2	US-08-473-553A-6
14	103	3.2	985	2	US-08-473-553A-2
15	102	3.2	1257	3	US-09-081-149-7
16	102	3.2	582	4	US-08-526-964-3
17	102	3.2	777	1	US-08-946-617-3
18	102	3.2	777	2	US-09-031-897-3
19	102	3.2	777	3	US-08-910-731-2
20	101.5	3.2	456	2	US-08-910-731-2
21	101.5	3.2	456	2	US-08-795-395-2
22	101.5	3.2	1817	4	US-09-004-838-125
23	101	3.2	461	2	US-08-910-731-6
24	100.5	3.2	264	4	US-09-543-681A-8068
25	100.5	3.2	1323	4	US-09-004-838-90
26	100	3.1	1255	3	US-08-947-823-3
27	98.5	3.1	478	4	US-09-004-838-51

28	98.5	3.1	885	2	US-08-310-912A-2
29	98.5	3.1	885	3	US-08-841-089-2
30	98.5	3.1	885	3	US-09-301-085-2
31	98.5	3.1	885	5	PCT-US95-04570-2
32	98.5	3.1	885	5	PCT-US95-04589-2
33	98	3.1	409	4	US-09-352-991A-29674
34	98	3.1	922	4	US-09-328-352-5283
35	98	3.1	1206	4	US-09-245-928A-19
36	98	3.1	1243	4	US-09-245-928A-17
37	98	3.1	1247	4	US-09-803-286A-2
38	98	3.1	1257	4	US-09-245-928A-18
39	97.5	3.1	320	4	US-09-325-932A-190
40	97	3.1	456	2	US-08-910-731-8
41	97	3.1	777	4	US-09-257-770-7
42	97	3.1	1568	3	US-09-181-706-2
43	97	3.1	1568	3	US-09-458-791-2
44	97	3.1	1568	3	US-09-459-066-2
45	97	3.1	1568	4	US-09-459-065-2

## ALIGNMENTS

RESULT 1  
US-09-172-841-51  
; Sequence 51, Application US/09172841  
; Patent No. 6232081  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeffrey W.  
; APPLICANT: Elledge, Stephen J.  
; TITLE OF INVENTION: F-BOX PROTEINS AND GENES  
; FILE REFERENCE: BCM-03510  
; CURRENT APPLICATION NUMBER: US/09/172.841  
; EARLIER FILING DATE: 1998-10-15  
; EARLIER APPLICATION NUMBER: 08/951,621  
; EARLIER FILING DATE: 1997-10-16  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 51  
; LENGTH: 456  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-172-841-51

Query Match 5.2%; Score 164.5; DB 3; Length 456;  
Best Local Similarity 23.8%; Pred. No. 2e-09;  
Matches 76; Conservative 51; Mismatches 124; Indels 69; Gaps 12;

QY	233	NLVSVKLTDCETLDLVNF-----FKHASALEEFCGGTYNEEPER-----YSAISLPAKL	281
DB	45	NILALDGSNWQRIDLFNFQIDVEGRVVRNI SKRCGGFLRKULSLRGCI GVGDSSLKXTFAQN	104
QY	282	CB-----LGTITGKNELPVFVFAAVLKLDLYAMLDTEDHGM-LIQRCNPNEVLE	333
DB	105	CNRIEHLNNGCKTIDTCTVSLSPFCKLHLDLTSCVINSILKSGISEGRNLEYLN	164
QY	334	TR--NVIGDRGLEVLRCCCKRLKRLIERGDDDDQGMEDBEGTVSHRGLIALSQGSELEY	391
DB	165	LSWCDQITKDGLEALVRGCGKAL-----LLRGCTQLED	199
QY	392	MAVYSDITNASLEHIGTHLKNLCLDFRVLVDHE--EKITDLPDNGVRAALLRGCDKLAR	449
DB	200	EAL-----KHIQNYC-HELVSINLQSCSRIT-----DEGVVQICRGCHRLQA	240
QY	450	FALYLRGGLTDVGLGYIGQYSPNVWMLLGVVGE--SDAGLLEFAKGCPSLQKLENGCL	508
DB	241	LCL--SGCSNLTDSITAGLNCNPRQILAEARCSHLTDAGFTLLARNCHELEKMDLEECI	299
QY	509	FFSERALAAVAATQATSLRYL	528
DB	300	LITDSTLIQLSIHCPIKQAL	319



RESULT 2  
US-08-951-621-51  
; Sequence 51, Application US/08951621  
; Patent No. 6573094  
; GENERAL INFORMATION:  
; APPLICANT: HARPER, JEFFREY W.  
; APPLICANT: ELLEGE, STEPHEN J.  
; TITLE OF INVENTION: F-BOX GENES AND PROTEINS  
; NUMBER OF SEQUENCES: 59  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Medlen & Carroll, LLP  
; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/951,621  
; FILING DATE: 16-OCT-1997  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MacKnight, Kamrin T.  
; REGISTRATION NUMBER: 38,230  
; REFERENCE/DOCKET NUMBER: BCM-02999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 51:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 456 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: protein  
US-08-951-621-51

Query Match 5.2%; Score 164.5; DB 4; Length 456;  
Best Local Similarity 23.8%; Pred. No. 2e-09;  
Matches 76; Conservative 51; Mismatches 124; Indels 69; Gaps 12;  
QY 233 NLVSVKLTDCIEDLVNF-----FKHASALEEEFCGGTYNEEPER-----YSAISLPAKL 281  
DB 45 NILALDGSNWQRIDLNFQIDVGRVVENISKRCGGFLRLSLRGCGIGVGDSSLKTFQ 104  
QY 282 CR-----LGLTVIGKNELPIVFMFAAVLKLDLYAMLDTHCHW-LIQRCPNLEVL 333  
DB 105 CRNIEHLNNGCKTITDSTCYSLRSCSKLKHLDLTSCVITNSSLKGISEGCRNLEYLN 164  
QY 334 TR--NVIGDRGLEVLGRCCXRLKRLRIERGDDOGMEDEGTGSHRGLIALSQCCELEY 391  
DB 165 LSWCDQITKDGIALVRCRCGLKAL-----LIRGCTQLED 199  
QY 392 MAVTVSDITNASLEHIGTHLKNLDCDFELVLLDHE--EKITDPLDNGVRALLGCDKLR 449  
DB 200 EAL-----KHTONYC-HELVSUNLQSCSRIT-----DEGVQICRCHRLQA 240  
QY 450 FALYLRGGGLTDVGLGTYGQYSPNVRWMLLYGYGE--SDAGLLEPAKGCPSLQKLEMRGCL 508  
DB 241 LCL--SGCSNLTDSITALGNLCPRLQIILEARCSHLTDAGFTLLARNCHELEKMDLECI 299  
QY 509 FFSERALAVAATQLTSLRYL 528  
DB 300 LITDSTLIQLSIHCPKLOAL 319

RESULT 3  
US-09-177-165A-31  
; Sequence 31, Application US/09177165A

Patent No. 6426205  
; GENERAL INFORMATION:  
; APPLICANT: Tyers, Mike  
; APPLICANT: Willem, Andrew  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING UBIQUITIN  
; TITLE OF INVENTION: DEPENDENT PROTEOLYSIS  
; FILE REFERENCE: 11757.10USUI  
; CURRENT APPLICATION NUMBER: US/09/177,165A  
; CURRENT FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: 60/092,443  
; PRIOR FILING DATE: 1998-07-10  
; PRIOR APPLICATION NUMBER: 60/063,254  
; PRIOR FILING DATE: 1997-10-24  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 31  
; LENGTH: 1151  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae  
US-09-177-165A-31

Query Match 4.0%; Score 127.5; DB 4; Length 1151;  
Best Local Similarity 20.7%; Pred. No. 0.00018;  
Matches 119; Conservative 78; Mismatches 190; Indels 189; Gaps 29;  
QY 11 LSLRSNMTEERNVRKTRVVVDVDCVIPY-----IDDPKDRDAV----- 49  
DB 212 MELRALETENMEIRNLRL--KILTIIEYKKSLEYACHSKLGGQQVNFDTNFIIWINSI 269  
QY 50 -----SQVCRWVELDSLTKHVTIAL-----CYTTTPARLRRRFP----- 85  
DB 270 DTTESSDLKEGLQSLRSYRQRFINNVLSNPNSQNTCTSVT-----RRSPFALNMLPSEI 324  
QY 86 -HLESILKLGKPRAMFNLPEDWGGHVTTPWVKEISQYEDCLKSLHFRMIKVSQDLQNL 144  
DB 325 LHLILDKNQYDIVKFLTVSKLWA-----EI-----IVKILYRPHINKKSQDLDF 371  
QY 145 ARDRGHVHLAKLKDCKSGTTDGLPHIGRPFCKSLRVLFLEESSILEKDGELWELALNNT 204  
DB 372 -----LRTMKLT-----SEETVFNYRLMIKRLNFSFV-----GDYHDTLNYF 410  
QY 205 VLETLNFIYTDIAVVKIEDELLAKNCPNLVSKLTDCEILDVNFPPKASALEEFCGGT 264  
DB 411 V-----GCKNL-----ERLTIV-FCKHITSV----- 430  
QY 265 YNEEPERYSAISLPAKLRLGLTYIGKNELPIVFMFAAVLKLDLYAMLDTHCHW--L 322  
DB 431 -----PISAVLRGCKF-----LQSVDTIT-GIRDVSDVDVFDTL 461  
QY 323 IQRCPNLE--VLETRNVIGD--RGLVLRGCCXRLKRLRIERGDDOGMEDEGTGSHR 377  
DB 462 ATYCPRVQGYFVQARNVTDFSLRNFVHS-----PMLKRIKITANNR-----MND 508  
QY 378 GLI-ALSQCCELEYNAVYVS-DITNASLEHIGTHLKNLDCDFELVLLDHEEKITDLELDN 435  
DB 509 -LVELLANKCPLLVEVDITLSPNVTSSLLKULTRVLQREFRIT---HMTNIT---DN 560  
QY 436 GVRALLRGCDKLRFPALYLRG--GLTDVGLGTYGQYSPNVRWMLLYGYGE--SDAGLLEF 492  
DB 561 LFQELSKVVDMPSLRLIDLSCGENTDKTIESIVNLAFLKRLNVLFGKCSRITDASLFOL 620  
QY 493 AKCPSLQKLEMRGCLFFSERALAVAATQLTSLRYL 528  
DB 621 SKLGNLQTVFHGCFNITDNGVRALFHSCTRIOYV 656

RESULT 4  
US-09-904-615-154  
; Sequence 154, Application US/09904615  
; Patent No. 6566325  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 49 Human Secreted Proteins

```

; FILE REFERENCE: P2032P1
; CURRENT APPLICATION NUMBER: US/09/904,615
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 154
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (194)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (205)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-904-615-154

Query Match      3.7%; Score 119; DB 4; Length 483;
Best Local Similarity 21.7%; Pred. No. 0.00036;
Matches 96; Conservative 56; Mismatches 150; Indels 140; Gaps 19;

QY 90 LKLGKPKRAAMNFI-IPEDW-----GGHVTWPVWKEISQYFDCLKSLHFR 132
Db 111 LSLVGRP-AGLSLRVQEPWADRARVLSLELVECAQASGVTEISITSETQQQLCVQ-LEFP 168
QY 133 RMIVKDSLDQNLARDRGHVLHAKLDKCSGTTDGLF--HIGRFCKSLRVLFLESSILE 190
Db 169 R-----QENPEAVARLHACDLGAHSHSLXGQMTCARLXQLSQVNICE 216
QY 191 KDGE---WLHELALNNTVLETFNLYLTDIAVVKIEDLELLAK---NCPNLVSVKLTDCI 244
Db 217 DDDASSLLQSLLSLELKT--FLTS-SCVSTEGLAHLASGLCHHL-----EE 265
QY 245 LDLVNFFKHASALEBFCCGTYNNEPERYSALSLPAKLCRLGLTYGKNELPIVFMFAVL 304
Db 266 LDLSN-----NQDEEGTKALMRALEGKW-----ML 291
QY 305 KKLDLLYAMLDTEDHOMLIQPCNPVLEET-----RNVDGRGLEVLGRCKELKELRIER 360
Db 292 KRLDLSHLLNSSTLALLTHRLSQMTCLQSLRLNRSIGDVG-----CCHLSEALR--- 342
QY 361 GDDQGMEDDEGTSHRGLIALSQGCSELEYMAVYVSDITNASLHIGTHLKNLCDFRLV 420
Db 343 -----RATSELDLSHNQIGDAGVQHILATILPGLPELRKI 378
QY 421 LLDHE-----EKITDPLDNGVRALLRGCDKRLRFALYLRRGGLTDVGLGYIGQYSP 472
Db 379 DLSGNSISSAGGVQLAESLVLCRLLEMLGCNAL-----GDPTALGLAQ--ELPQ 427
QY 473 NVRWMLLVGVSDESAGLLLEFAK 494
Db 428 HLRVLHLPFSLHGPFGALSLAR 449

RESULT 5
US-09-172-841-53
; Sequence 53, Application US/09172841
; Patent No. 6232081
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeffrey W.
; APPLICANT: Elledge, Stephen J.
; TITLE OF INVENTION: F-BOX PROTEINS AND GENES
; FILE REFERENCE: BCM-03510
; CURRENT APPLICATION NUMBER: US/09/172,841
; CURRENT FILING DATE: 1998-10-15
; EARLIER APPLICATION NUMBER: 08/951,621

```

```

; EARLIER FILING DATE: 1997-10-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 53
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-172-841-53

Query Match      3.6%; Score 113.5; DB 3; Length 380;
Best Local Similarity 20.7%; Pred. No. 0.00098;
Matches 97; Conservative 65; Mismatches 131; Indels 175; Gaps 24;

QY 57 YELDSLRKHVTTA-ICVTTTPARLR-----RPHLESKLKPKRAAMNLIPEDW 108
Db 48 YELIQLILNHLTPDLCLRLAQCTCKLSHQHCCDPLQYIHL-----NLQY-- 91
QY 109 GGHVTPVWKEISQYFDCLKSLHFRFMIVKDSLDQNLARDRGHVLHAKLDKCSGTTDGL 168
Db 92 -----WAK-----LDDTSLEFL-QSRCTLVQWNLN---SWTGNRGF 123
QY 169 FHIGRFCKSLRVLFLESSILEKDGWELHELALNNTVLETFNLYLTDIAVVKIEDLELLA 228
Db 124 ISVAGFSRFLKVCSELVEL-----ELSCSHFLNETC-----LEVIS 160
QY 229 KNCNPLVSVKLTDCIHLDLNFPFKHASALEBFCCGTYNNEPERYSALSLPAKLC---RLG 285
Db 161 EMCNQLQALNLSSCDKL-----PPQAFNHI---AKGCSLKKLV 195
QY 286 L--TYIGKNELPIVFMFAVLKLDLLYAMLDTEDHOMLIQPCNPVLETRNVIGRGL 343
Db 196 LYRTKVSQTALLSILNFCSELQHLSL-----GSCVMI-----EDYDVA--- 234
QY 344 EVLGRCKRLKRLRIERGGDDDDQGMEDBEGVSHRGLIALSQGCSELEYMAVYVSDITNAS 403
Db 235 SMIGAKCKKLTLDLWRCKN-----ITENGIAELASGCPLEELDLGWCPTLOSS 284
QY 404 -----LEHIGTHLKNL-----CD-----FRVLDDHEEKITDPLDNGVRA 439
Db 285 TGCFTRLAHQPLNQLFLTANRSVCDTDIDELACNCTRLQQLDLGKVTI-----YKF 338
QY 440 LIRGC--DKLRRFALYLRRGGLTDVGLGYIGQYSPNVRWMLLVGVGES 485
Db 339 VLNVCFDDRKANLRLVVRKKIFGVKNFI-----LIRW--LGLIGNA 379

RESULT 6
US-08-951-621-53
; Sequence 53, Application US/08951621
; Patent No. 6573094
; GENERAL INFORMATION:
; APPLICANT: HARPER, JEFFREY W.
; APPLICANT: ELLEDGE, STEPHEN J.
; TITLE OF INVENTION: F-BOX GENES AND PROTEINS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/951,621
; FILING DATE: 16-OCT-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: MacKnight, Kamrin T.

```

REGISTRATION NUMBER: 38,230  
REFERENCE/DOCKET NUMBER: BCM-02999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 705-8410  
TELEFAX: (415) 397-8338  
INFORMATION FOR SEQ ID NO: 53:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 380 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: protein  
US-08-951-621-53

Query Match 3.6%; Score 113.5; DB 4; Length 380;  
Best Local Similarity 20.7%; Pred. No. 0.0098;  
Matches 97; Conservative 65; Mismatches 131; Indels 175; Gaps 24;

QY 57 YELDSLTRKHVTIA-LCYTTTPARLR-----RPHLESKLKGPRAAMFNLPEDW 108  
DB 48 YELIQLIINHLTPDLCLIAQCKLSQHCCDPLQTIHL-----NLQY-- 91  
QY 109 GGHVTPWVKETISQYFDCLSLHFRMRIVKDSLOQLARDRGHVLHALKDKCGFTDGL 168  
DB 92 -----WAK-----LDDTSLEFL-QSRCTLVQWNLN---SWTGNRGF 123  
QY 169 PHIGPCKSLRVLFESSILEKGEWLHLMNNTVLTNPFYLTDAVVKIEDLELLA 228  
DB 124 ISVAGFSRFLKVCSELVRL-----ELSCSHFLNETC-----LEVIS 160  
QY 229 KNCPLNVSVKLDCETLIDLVNFFKHASALEEFCGGTYNEPEPERYSALSLPAKLC--RLG 285  
DB 161 EMCNPLQALNLSCKL-----PPOAFNHI-----AKLSLKLRLV 195  
QY 286 L--TYIGKNELPIVFMFAAVLKKLDLYAMLTDEHMLIQRCPNLEVTETRVIGDRL 343  
DB 196 LYRTKVEQFALLSILNFCSELQHLSL-----GSCVMI-----EDYDVIA--- 234  
QY 344 EVLGRCKKELRLRIERGDDQDQMEDEEVSVHRGLIALSQGSELEYMAVYVSDITNAS 403  
DB 235 SMIGAKCKLRTLDLWRCKN-----ITENGIAELASGCPLELDLGLWCPTLQSS 284  
QY 404 -----LEHIGTHLKNL-----CD-----FRLVLDHEKITDPLONGVRA 439  
DB 285 TGCFTFLAHLQNLQKLFITANRSVCDTIDELACNCTRLQQLDILIGKVTI-----YKF 338  
QY 440 LLRGC--DKLRFALYLRGGLTVDGLGIVGYQVSPNRMMLGYGES 485  
DB 339 VLNVCPLDRKANLRLFVRKKKIFGYNKFI-----LIRW--LGLIGNA 379

RESULT 7  
US-08-531-439B-4  
Sequence 4, Application US/08531439B  
Patent No. 5981702  
GENERAL INFORMATION:  
APPLICANT: Zhang, Hui  
APPLICANT: Beach, David  
TITLE OF INVENTION: Cyclin/CDK Associated Proteins,  
TITLE OF INVENTION: and Uses Related Thereto  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley, Hoag & Eliot  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII (text)

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/531.439B  
FILING DATE: 21-SEP-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: CSV-006.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 832-1000  
TELEFAX: (617) 832-7000  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 435 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-531-439B-4

Query Match 3.5%; Score 112.5; DB 2; Length 435;  
Best Local Similarity 23.2%; Pred. No. 0.0016;  
Matches 72; Conservative 34; Mismatches 75; Indels 129; Gaps 17;

QY 49 VSOVCRRWYEL---DSLTRKHVTIALCYTTTPARLRPPHLESKLKGPRAAMFNLP 105  
DB 132 VSGVCKRWYLASDESL-----WQTLDTGK-----NLHP 161  
QY 106 EDWGHVTPWVKETISQ---YFDCLKSL-----HFRMRIVKDSLOQLARDRGHVLHAL 155  
DB 162 DVTG-----RLLSQGVIAFRCPSPFMDQPLAEHFSFRVQMDLSNVIEWS-TLHGI 213  
QY 156 KLDKCSGFTTDLGPHIGRFSKLSRVLFLEESSILEKGEWLHLMNNTVLTNPFYLT 215  
DB 214 -LSQCSK-----LQNLSL-----ELRLSDPIVNT--- 236  
QY 216 IAVKIEDLELLAKCNPVSVKLTDC---EILDVNFKHASALEE---FCGGTYNEE 268  
DB 237 -----LAKN-SNLVRINLPGCFKFPKPLQTLSSCPRLDELNLNLSWC-PNFTEK 283  
QY 269 PERYSALSLPAKLCRLGTYIGKNELPIVFMFAAVLKKLDLYAMLTDEHMLIQRCPN 328  
DB 284 HVQVAHVSETWQNLNSGYKN-----LQKSDL-----STLVRRCPN 322  
QY 329 LEVLETRNVI 338  
DB 323 LVHLDLSNSV 332

RESULT 8  
US-09-331-403-2  
Sequence 2, Application US/09331403  
Patent No. 6489147  
GENERAL INFORMATION:  
APPLICANT: ALTMANN-JOHL, Regula; PHILIPPSEN, Peter; ALTHOFER,  
APPLICANT: Henning; SEULBERGER, Harald.  
TITLE OF INVENTION: Adenylate cyclase gene, and its use  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Keil & Weinkauff  
STREET: 1101 Connecticut Avenue  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage  
COMPUTER: IBM AT-compatible, Pentium III processor  
OPERATING SYSTEM: Windows 95  
SOFTWARE: WordPerfect version 6.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/331,403  
FILING DATE: 21-Jun-1999  
CLASSIFICATION: <Unknown>

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP97/07309
; FILING DATE: 29-DEC-1997
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1874 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-331-403-2

Query Match          3.5%; Score 111.5; DB 4; Length 1874;
Best Local Similarity 22.7%; Pred. No. 0.029;
Matches 107; Conservative 71; Mismatches 160; Indels 133; Gaps 27;

QY 79 RLRRFPHLSLKLKGPRAAMFNLIPEDDGHHVTPWVKESIQYFDCILKSLHFRMIVKO 138
Db 721 RLPRGFQDLKSLQL--DISSKNFNIYEVINSCN--LLQLDLSYNKIRSL-----PD 770
QY 139 S--DLQNLAR--DRGHVHIAKLDKCGFTTDLGLFHIGRF-----CK--SLRVLFLE 184
Db 771 SMNQLOKLAKINLSNNRITHVNDLSKMTSLRTDL---RYNRIESIKRVPNQLNFLT 826
QY 185 ES--SILEKGEWHLALANNVLETL---NFYLTDAIVVKIEDLEL-----LAKNCPN 233
Db 827 ENRLTMFDDQLMLRTLELQRPNSILTKNDYLEHLTSLSKAKUAVLPESLRLRPR 886
QY 234 LVSVKLTDCETILDVNFFFKHASALEEFCGGTYNEE--PERSAI-----SL 277
Db 887 LEKLELSNSITVLPDPIKHKLKLVHLSVAKNKLESDEIASLKNLAKMLDHCNNMLT 946
QY 278 PAKLCRLGLTYI-----GKNEL-----PIVFMFAA-----302
Db 947 PAALSTLSITFNISSNMLSGHHLYRTFQGTSTIAKSLMFLSAADNQMGDKFWEIFNTF 1006
QY 303 -VLKKLDLYAMLTDEHCLMIOQPNLEV--LETRNVIGRGLEVLGRCCCKRLKRLAI- 358
Db 1007 KTLKVLNLSY-----NNFMAL---PELEMENLTSLYSGNHLTTLTSGEAFIKLSLRLV 1057
QY 359 -ERGGDDGMEDEGTVSHRGIALSQGCESELEY-MAVVVSDIT---NASLEHI-----407
Db 1058 MLNANNLQSLPAEISQLSQ--LSVIDVGSNQLKXISNVHYDNNRQNTKYLNFSGNT 1115
QY 408 -----GT--HLKNLCDFR-----LVLLDHEEKITDLPIDNGVRLLR 442
Db 1116 RPEIKSAMEYGTNNMHSLSLTVLKQLRVLGLMDVTLNLSRVP--DDGVNFRLR 1165

RESULT 9
US-09-325-932A-185
; Sequence 185, Application US/09325932A
; Patent No. 6451604
; GENERAL INFORMATION:
; APPLICANT: Flinn, Barry
; TITLE OF INVENTION: Compositions affecting programmed cell
; TITLE OF INVENTION: death and their use in the modification of forestry plant develop
; FILE REFERENCE: 1022
; CURRENT APPLICATION NUMBER: US/09/325,932A
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 185
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-325-932A-185

Query Match          3.5%; Score 110.5; DB 4; Length 579;
Best Local Similarity 22.0%; Pred. No. 0.0046;
Matches 108; Conservative 70; Mismatches 158; Indels 155; Gaps 30;

```

```

QY 152 LHALKLDKCSGFTTDLGLFHIGRF-----CKSLRVLFLEESILE 190
Db 73 LSLLEILLRGESFVESIFSAGEIRQLVYLQWKECPISISFTTIPTRNLVLYIQYALKT 132
QY 191 KGEWHLALANNVLETLNFY-----LTDI--AVVKIEDLELA-----KNCPN-----233
Db 133 L---WQHE---SQAPLQLTLYIDATLSRVFQSIGKLNQLERIVLKNKYFKTLFNEFYDM 186
QY 234 --LVSVKLTDCETILDVNFFFKHASALEEFCGGTYNEEPERSAIS-----LPKLCRL--GL 286
Db 187 HSLKHITLQCEQOML-----LPDSVGILTGROTHDFSGCSNLOALPDSVGQLTGL 237
QY 287 TYTGKNE-----LPVFMFAAVLKLDLYAMLTDEHCLMIOQPNLEVLETRNVIGD 340
Db 238 KTLDELECTSLQGLPDSVGQLTGLQSLDL-----EHCSTLQGLP-----DSVGQ 281
QY 341 -RGLEVIG-RCCKKLKKLRIRERGGDDGMEDEGTVSHRGIAL-SOGCSLEYMAVVVS 397
Db 282 LTGLQTLDLRGCSL-----QGLPDSVGQLT--GLEGLYLSGCFSLQGLPDSVE 328
QY 398 DITNA-----SLRHI-----GTHLKNL--CDFRLVLLDHEEKITDLPIDNGVR 438
Db 329 QLTGLEGLYLSGCFSLQGLPDSVGQLTGLQSLNLEYCTSLGLEPDSVGQLTDLPLD 385
QY 439 ALLRGCDKLRRFALYLRGGGLTDVGLGYI--QOYSPNVRMMLLVYVGSDDAGLEFAKGP 497
Db 386 --INTCISLQ-----GLPD--SVGQLRGLQNLRLWC-----DSLQGLPDSVGQLT 427
QY 498 SLOKLEWRGCLFFGERALAVAAQTLSIRYLVWOG-----YGVSPSGRDLVWARPFW---550
Db 428 GLQILDLGGCT--SLQGLPDSVGQLTGLRTLHLENCTSLQGLPDSVGNLTSIK---WLNL 482
QY 551 ----NIELIPS 557
Db 483 SGCSNLOMLPN 493

RESULT 10
US-08-462-467B-12
; Sequence 12, Application US/08462467B
; Patent No. 6210899
; GENERAL INFORMATION:
; APPLICANT: Rosenbaum, Jan S
; TITLE OF INVENTION: The Use of a BMP Protein Receptor
; TITLE OF INVENTION: Complex for Screening Bone Metabolism Actives and Cells
; TITLE OF INVENTION: Co-Transfected With a Type II BMP Receptor and a Type I
; TITLE OF INVENTION: BMP Receptor
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Procter & Gamble Company
; STREET: 11810 East Miami River Road
; CITY: Ross
; STATE: OH
; COUNTRY: USA
; ZIP: 45061
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,467B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Herisko, Bart S.
; REGISTRATION NUMBER: 32,572
; REFERENCE/DOCKET NUMBER: 5474R
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (513) 627-0633
; TELEFAX: (513) 627-0260
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:

```

LENGTH: 1495 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: not relevant  
 TOPOLOGY: not relevant  
 MOLECULE TYPE: protein  
 US-08-462-467B-12

Query Match 3.4%; Score 109; DB 3; Length 1495;

Best Local Similarity 22.3%; Pred. No. 0.038;  
 Matches 107; Conservative 44; Mismatches 164; Indels 164; Gaps 20;

QY 167 GLFHGRCKSLRVFLFESSILEKDGWHLHNLNNTVLETLNFIYLDI-----AVVKI 221  
 Db 772 GLTHRGILLETYRGLNTHRVALLMETARGHISGLASNILEGLYPHEILEALAAAS 831  
 QY 222 EDLELLA-----KNCP-----NLVSKLTCEDILVNPFKHASALEEFCGGTY 265  
 Db 832 PILELYSGLYTHRGYSERTRTHRGLENTYRLEILETHRASPTYRHISGLAS--NGLY 889  
 QY 266 NEPEPERSAISLPKLCRLGLTYIGKNELPIVFMFAVLKDLIYAMLDTHEDHMLIOR 325  
 Db 890 SERLETYRA-SPHEL-----ELYSYCSALATH-----916  
 QY 326 CNLNVLETRNVIGRGLEVLGRCKRLKRLRIERGDDDDQGDDEGTVSHRGLIALSOG 385  
 Db 917 --RLEASPTRARGALALELYSLEALATYR-----SERALAAAL--955  
 QY 386 CSELEYMAVYVSDITNASLEHIGTHLKNLCDFRLVLDHEEKITDPLDNGVRALLRGCD 445  
 Db 956 C-----YSGHLYBCYSHLEHISTH-----976  
 QY 446 KLRREYALYRRGLTVDVGLYIGQYSPNVRMML--LGYVGBSDAGLLEF-----AKG 495  
 Db 977 --RGLLETYRGLYTHRGNGLYSPRALILEALAHISARGASPLELYSSERLYSASN 1034  
 QY 496 CPSLOKLEWRGCL-----FFSERAL-----AVATQTLSTLYLWVQGVV-----535  
 Db 1035 ILELEILELYSLSYASNGLYSERCYSCYSILEALASPLEGL--YLEALAVALLYSPEAS 1093  
 QY 536 ----SPSGR--DLVMPARPPWNIELIPSRKVATNTNPDVTVVVHPAHILAYSLAG 586  
 Db 1094 NSERASPTHRASGLVALASP-----ILPRLEASNTHRARGVALGLYTHRU--YSARG 1145

RESULT 11

US-09-252-991A-28407  
 ; Sequence 28407, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 28407  
 ; LENGTH: 870  
 ; TYPE: PRT  
 ; ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-28407

Query Match 3.4%; Score 108; DB 4; Length 870;  
 Best Local Similarity 21.6%; Pred. No. 0.019;  
 Matches 140; Conservative 79; Mismatches 240; Indels 190; Gaps 35;

QY 77 PARLRRRFFHL-----ESLKLKGPRAAMPNLIPEDMGCHVTYTPWKETISQYFDCILKS 128  
 Db 88 PATFRRLAGALGPESDEAQLCRRLNASALWLLDDYPRHPDP---AFDACLDCLLS 144

QY 129 -----LHFRMIVKSDILO-----NLARDGHVHLKLDKCSGFTTGG 167  
 Db 145 AASPRVGVWLASRRPQCINLARLLLEGELLEVGGGLAFDADEVAELRL--HGRSADA 201  
 QY 168 -----LFHIGRFCKSLRVFLF--EESILEKDGWHLHNLNNTVLETLNFIYLDI 219  
 Db 202 AATGALLERSGGWCNARLRLLVGEETA-----GPLLQYV-LQHELLDELPPPLADAARA 255  
 QY 220 -----KIBDL-----ELLAKNCPNLVSVK-----238  
 Db 256 LAMPQVGPPELFRRLFEELPHGLDILLARGPPLADANQRYSLPPAIRREARPPQR 315  
 QY 239 ---LTDCEILDVNFFKHASALEEFCGGTYNEEPERSAISLPKLC-----RLGIT 287  
 Db 316 DFHREACEW-----FASGETREAVDQALAADEP--TAAGLLQKLTTEEQLLHGHNI 368  
 QY 288 YIGKNELPIVFMFA-----AVLKLDLLYA--MLDTEHCH--LIQRCNLEVLETRN 339  
 Db 369 LALRDELPAALLASTPRVLNLTWLTLYAGRLAEAD-CIGLARFLPMPASRQVILLA 427  
 QY 340 D-RGI-EVLGRCCRLKRLRIERGDDDD-----QGMED--EETVSHRGLIALSOG 388  
 Db 428 QWQGLFGILLHC-----RGERGAADYLRLEALQEPEDAWSQGLICRSALMQLAI 480  
 QY 389 LEYMAVYVSDITNASLEHIGTHLKNLCDF-RLVLDHEEKI--TDLPLDNGVRAL--440  
 Db 481 MDQARLIGDALRLAREHDSLIFEALIELEAQLMQRGELLRAEGVLDRAQRYLEDL 540  
 QY 441 ---LRGCDKLRFPALYLRGGLTVDVGLY---IQYSPNV-RMMLGYVGBSDAGLLEF 492  
 Db 541 QGSPMLGRALARRARLCLOQGREVEBAGHWYRLGLEQARENLDPMALYGLG--LALLEA 597  
 QY 493 AKG-----CPSLOKLE-----MRGLFFSERALAAVAATQTLSTLYLWVQGV 535  
 Db 598 GQGLDAAAFNRILLEVERLWQRHVPDPLRYRGALLVSSALTLO-----QGR 643  
 QY 536 SPSSGDLVMPARPPWNIELIPSR-KVATNTNPDVTVVVHPAHILAYYS 583  
 Db 644 PAQAREILLRVRAYFQ---PGRARLSPPEPELEARVHQALALAEYS 688

RESULT 12

US-09-004-838-111  
 ; Sequence 111, Application US/09004838  
 ; Patent No. 6350933  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Michelmore, Richard W.  
 ; APPLICANT: Shen, Kathy  
 ; APPLICANT: Meyers, Blake  
 ; TITLE OF INVENTION: Procedures and Materials for  
 ; TITLE OF INVENTION: Conferring Pest Resistance in Plants  
 ; NUMBER OF SEQUENCES: 140  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Townsend and Townsend and Crew LLP  
 ; STREET: Two Embarcadero Center, Eighth Floor  
 ; CITY: San Francisco  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 94111-3834  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/004,838  
 ; FILING DATE: 09-JAN-1998  
 ; CLASSIFICATION: 800  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/781,734  
 ; FILING DATE: 10-JAN-1997  
 ; ATTORNEY/AGENT INFORMATION:



349 QY CKK-----RLKRLIERGD--DDQMEDE---EGTVSHRGLIALSQSCSELEYMAVYVS 397  
 714 Db GMPNNVDVAIKRL-VORGTGRSDHGTAEIQTGLRHRHIVRLGVVANKDTNLLYE 772  
 398 QY DITNASLEHI-----GTHL-----KNLC-----DFRLVLDHEKITDPLDN 435  
 773 Db YMPNGSLGELLHGKGGHQLQWETRRVAVEAAKGLCYLHDCSPILHRDVKNNNILDS 832  
 436 QY GVRA 439  
 833 Db DFEA 836

RESULT 15

US-08-947-823-5  
 ; Sequence 5, Application US/08947823  
 ; Patent No. 6114605  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Williamson, Valerie M.  
 ; APPLICANT: Kaloehian, Isagouhi  
 ; APPLICANT: Yaghoobi, Jafar  
 ; APPLICANT: Bodeau, John  
 ; APPLICANT: Milligan, Stephen  
 ; TITLE OF INVENTION: Procedures and Materials for Confering  
 ; TITLE OF INVENTION: Pest Resistance in Plants  
 ; NUMBER OF SEQUENCES: 5  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Townsend and Crew LLP  
 ; STREET: Two Embarcadero Center, Eighth Floor  
 ; CITY: San Francisco  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 94111-3834  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA: US/08/947,823  
 ; FILING DATE: 09-OCT-1997  
 ; CLASSIFICATION: 800  
 ; PRIOR APPLICATION NUMBER: PCT/US97/18802  
 ; FILING DATE: 09-OCT-1997  
 ; PRIOR APPLICATION DATA:  
 ; FILING DATE: 10-OCT-1996  
 ; APPLICATION NUMBER: US 60/028,191  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Bastian, Kevin L.  
 ; REGISTRATION NUMBER: 34,774  
 ; REFERENCE/DOCKET NUMBER: 023070-070210US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 576-0200  
 ; TELEFAX: (415) 576-0300  
 ; INFORMATION FOR SEQ ID NO: 5:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1257 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-947-823-5

Query Match 3.2%; Score 103; DB 3; Length 1257;

Best Local Similarity 21.0%; Pred No. 0.13;  
 Matches 107; Conservative 67; Mismatches 176; Indels 160; Gaps 25;

7 FLFTLSLRNMTTEENVRKRVVDVLDVPIYDDPKDRDAVSQCRRWYEL-DSLTFK 65  
 788 YLFTVYLG-EGFVEKTEKMGIE-EVVKIYMDLISLIV--IC--FNEIGDILNFQ 838  
 66 HVTIALCYTTTTPARLRPPHLESKLKGPRAAMENLIPEDWGGHVTWPWKEISQYFDC 125

US-08-473-553A-2  
 ; Sequence 2, Application US/08473553A  
 ; Patent No. 5859338  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Meyerowitz, Elliot M.  
 ; APPLICANT: Clark, Steven E.  
 ; APPLICANT: Williams, Robert W.  
 ; TITLE OF INVENTION: Plant Clavatal Nucleic Acids,  
 ; TITLE OF INVENTION: Transformed Plants, and Proteins  
 ; NUMBER OF SEQUENCES: 11  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Flehr, Hohnbach, Test, Albritton & Herbert  
 ; STREET: Four Embarcadero Center, Suite 3400  
 ; CITY: San Francisco  
 ; STATE: California  
 ; COUNTRY: United States  
 ; ZIP: 94111-4187  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/473,553A  
 ; FILING DATE: 06-JUN-1995  
 ; CLASSIFICATION: 800  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Silva, Robin M.  
 ; REGISTRATION NUMBER: 38,304  
 ; REFERENCE/DOCKET NUMBER: A-60886/RFT/RMS  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 781-1989  
 ; TELEFAX: (415) 398-3249  
 ; TELEX: 910 277299  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 985 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-473-553A-2

Query Match 3.2%; Score 103; DB 2; Length 985;  
 Best Local Similarity 19.9%; Pred. No. 0.086;  
 Matches 108; Conservative 71; Mismatches 191; Indels 174; Gaps 26;

9 FTLSLRNMTEENVRKRVVDVLDVPIYDDPKDRDAVSQCRRWYELSLTRKH-- 66  
 354 FTLQPLANLGRNGLIKLDVSDNHLTGLI-----PKD-----LC-RGEKLEMLLSNFF 401  
 67 -----VIALCYTTTTPARLRPPHLESKLKGPRAAMENL----- 103  
 402 FFGPIPEELGCKSKTKIRIVKNL-----LNGTVPAGLNLPLVTLIELTDNFFSGEL 454  
 104 -----IPEDW-GGHVTPWVKEISQYFCLKSLHFRMRVIVKSDIQLNLABRG 149  
 455 PVTMSGDVLQIYLSNNWFSGEIPPAIGN-----FENLOTFLDRNRFRG-----NIPRIF 506  
 150 HVLHAKLDCSGFTTDLGFLHGRCKSLRVLFLEBSILEKDGWELHELANNVTLETL 209  
 507 ELKHLSRINTSANNITGGIPDSISRCSTLISVDLSNRRI-----NGEIPKGI 553  
 210 NFVLTDAIVVKEDLEL-----LAKNCNPLVSVKLTDCIEILDVN-----FFKHASAL 257  
 554 N-NVKNLGILNIGNQLTGSIPITGIGNMTSLTLDLSFNDLSGRVPLGQFLVFNETS-- 610  
 258 EEFCCGTYNEEPERYSAISLPAKL-----CRGLGTYIG----- 290  
 611 --FAGNTYLCPLHRVSCPTPGQTSDBNHTALFSPSRIVITVIAATIGLILSVATROWN 668  
 291 --KNEPLVPMFAVLKLLDLLVAMLDTEDHCLMIQRCNPLEVLETRNVIGDRGLVLR 348  
 669 KKNQKSLAWKLA-FQKLD-----FKSED-----VLECLKEENIIKGGSGIVYR 713

Db 839 IHDLVHDFCUIKAKENLFRIRK-----SAPDLLPR-----QITIDYDEEEHF-- 884  
Qy 126 LKSLHFRMRIVKDSDLQNLARDRGHVLHAKL--DKCSGETTDGLPHIGRCKSLRVLFL 183  
Db 885 --GLNF---VMFDS---NKRHSGKHLYSLRINGDQLDDSVSDA-FHL-RHLRLIRVLDL 934  
Qy 184 EESSILEKDGEMWHELALNN----- 203  
Db 935 EPSLIMVND-SLLNEICMLNHLRIRIQVKYLPFPFSNLMNLESFVSNKGSILVLLP 993  
Qy 204 TVLETNLFYLDIAVVKIEDLE-----LLAKNCPNLVSKLTDCEIL-----DLVN 249  
Db 994 RILDVLKRLVLSVGACSFDDADESILIAKD-----TKLENRILGELLISYSKDTMN 1047  
Qy 250 FFKHASALEEFC-----GGTYNEBPERYSALPAKLCRLGLTYIGN-----EL 294  
Db 1048 IFKFPNLQVLQFELKESWDYSTQHWFPKLDCLTELETCVGFKSSNTNHCSSVVVTR 1107  
Qy 295 PIVMFAAVLKLDLILYAMLTEDHCLIQPCNLEVLTEENVIGDRGLEVLGRCKKELK 354  
Db 1108 PWFHFPNKLKEL-LLYDFPLTSDSLSTIARLPNLENLSLYDTI----- 1150  
Qy 355 RLRTERGDDOGMED-----BEGTVSHRGLIALS-QGCSELE----- 390  
Db 1151 ---IQGEWNNGEEDTFENLXFLNRLTLTKWVGEESFPNLEKXLOECGKLEEIIPS 1207  
Qy 391 ---YMAYVVSIDITNASLEHIGTHLKNLCD 416  
Db 1208 FGDIYSLAFIKIVKSPQEDSALKIKKYAE 1237

Search completed: April 20, 2004, 01:39:25  
Job time : 28 secs





Db	121	IVKSDQLQNRDRGHVHLKDKSCFTTDLGPHIGRCKSLRVLFLEESSILEKDCG	180
Qy	195	WLHELALANTVLETINFYLTDAVVKIEDLELLAKNCPLVSVKLTDCCEILDVNFPPKHA	254
Db	181	WLHELALANTVLETINFYLTDAVVKIEDLELLAKNCPLVSVKLTDCCEILDVNFPPKHA	240
Qy	255	SALEBFCGGTYNEEPKYSALSPAKLCRLGLTYIGKNELPIVFMFAAVLAKKLDLLYAML	314
Db	241	SALEBFCGGTYNEEPKYSALSPAKLCRLGLTYIGKNELPIVFMFAAVLAKKLDLLYAML	300
Qy	315	DTEDHMLIQRCPNLEVLTRNVI GDRGLEVLGRCKRLKRLRIERGDDDDQMEDEEGTV	374
Db	301	DTEDHMLIQRCPNLEVLTRNVI GDRGLEVLGRCKRLKRLRIERGDDDDQMEDEEGTV	360
Qy	375	SHRGLIALSQGCSELEYMAVYVSDITNASLEHIGTHLKNLCPDLRLVLDHBEKIIDPLD	434
Db	361	SHRGLIALSQGCSELEYMAVYVSDITNASLEHIGTHLKNLCPDLRLVLDHBEKIIDPLD	420
Qy	435	NGVRALLRGCDKLRFPALYLRGGTLDVGLGYIGQYSPNVRMMLGYVGS DAGLLEFAK	494
Db	421	NGVRALLRGCDKLRFPALYLRGGTLDVGLGYIGQYSPNVRMMLGYVGS DAGLLEFAK	480
Qy	495	GCPSLQKLEMRGCLFSEFALAAVAATQLTSLRYLWVQGYGVSFSGRDLVLPFNWIEL	554
Db	481	GCPSLQKLEMRGCLFSEFALAAVAATQLTSLRYLWVQGYGVSFSGRDLVLPFNWIEL	540
Qy	555	IPSRKVATNTNPDET VVVEHPAHILAYYSLAGORSDFPDVVPDTCVDT	606
Db	541	IPSRKVATNTNPDET VVVEHPAHILAYYSLAGORSDFPDVVPDTCVDT	592
RESULT 2			
US-10-424-593-284666			
; Sequence 284666, Application US/10424599			
; Publication No. US20040031072A1			
GENERAL INFORMATION:			
; APPLICANT: La Rosa Thomas J			
; APPLICANT: Kovacic David K			
; APPLICANT: Zhou Yihua			
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With			
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement			
; FILE REFERENCE: 38-21(53223)B			
; CURRENT APPLICATION NUMBER: US/10/424,599			
; CURRENT FILING DATE: 2003-04-28			
; NUMBER OF SEQ ID NOS: 285684			
; SEQ ID NO 284666			
; LENGTH: 590			
; TYPE: PRT			
; ORGANISM: Glycine max			
FEATURE:			
OTHER INFORMATION: Clone ID: PAT_MRT3847_99079C.1.pep			
US-10-424-599-284666			
Query Match 97.3%; Score 3095; DB 12; Length 590;			
Best Local Similarity 99.5%; Pred. No. 8.3e-306;			
Matches 587; Conservative 3; Mismatches 0; Indels 0; Gaps 0;			
Qy	17	MTEERNVKTRVVDVLDVCPYIDDPKDRDAVSQVCRWYELDSLTKKHVTTCALCYTTT	76
Db	1	MTEERNVKTRVVDVLDVCPYIDDPKDRDAVSQVCRWYELDSLTKKHVTTCALCYTTT	60
Qy	77	PARLRPPHLESKLKGPAAAFNLIPEDWGGHVPWKEISQYDCLKSLHFRMIV	136
Db	61	PARLRPPHLESKLKGPAAAFNLIPEDWGGHVPWKEISQYDCLKSLHFRMIV	120
Qy	137	KDSDQLNARDRGHVLHAKDKSCGFTTDLGPHIGRCKSLRVLFLEESSILEKDCGWL	196
Db	121	KDSDQLNARDRGHVLHAKDKSCGFTTDLGPHIGRCKSLRVLFLEESSILEKDCGWL	180
Qy	197	HELALNNTVLETINFYLTDAVVKIEDLELLAKNCPLVSVKLTDCCEILDVNFPPKHA	256

```

; APPLICANT: Zeng, Xiaoping
; APPLICANT: Zhang, Qiang
; APPLICANT: Zhao, Yajuan
; APPLICANT: Zhou, Li
; TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants
; FILE REFERENCE: 38-15(52796)B
; CURRENT APPLICATION NUMBER: US/10/310,154
; CURRENT FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: 60/337,358
; PRIOR FILING DATE: 2001-12-04
; NUMBER OF SEQ ID NOS: 736
; SEQ ID NO 520
; LENGTH: 598
; TYPE: PRT
; ORGANISM: Zea mays
; US-10-310-154-520

Query Match      56.2%; Score 1786.5; DB 15; Length 598;
Best Local Similarity 60.1%; Pred. No. 1.7e-172;
Matches 349; Conservative 85; Mismatches 136; Indels 11; Gaps 7;

QY 28 VDVVLDVIVYIDDPKRDVAVSOVCREWYELDSLTRKHVTIALCYTTTPARLRPPH 87
DB 22 VPEALHLVFGYDDPDREASVLCRUHRIKALSRKHVTGFCYAVEPARLLARPRL 81
QY 88 ESKLKGKPRAMFNLIPEDMGHHVTPWVKESIQYFDCLSLHFRMIVKDSLDQNLARD 147
DB 82 ESKLKGKPRAMFNLIPEDMGHHVTPWVKESIQYFDCLSLHFRMIVKDSLDQNLARD 141
QY 148 RGHVHLAKDKCSGFTTDLGPHGRFCKSRVLFLESSILEKDG-EWLHELALNTVL 207
DB 142 RGHVHLAKDKCSGFTTDLGPHGRFCKSRVLFLESSILEKDG-EWLHELALNTVL 201
QY 208 TLFVLTIDIAVKIE--DLELLAKNCPNLVSKLTDCEILDVNFPHKASALEEFCGGTY 265
DB 202 TLFVLTIDIAVKIE--DLELLAKNCPNLVSKLTDCEILDVNFPHKASALEEFCGGTY 258
QY 266 NE--EPERYSAISLPAKLCRL--GLTYIGKNELPIVFMFAAVLKKDLLEYAMLTDEHOML 322
DB 259 FEVGEYTKYKVIFFPRCLFGLTFMGKNEMPIFFYSTMLKKDLQFTLTTHDQCL 318
QY 323 IQCPNLVETRNVIQDGLVGRGLVGRGLVGRGLVGRGLVGRGLVGRGLVGRGLVGRGLV 382
DB 319 IAKCPNLVETRNVIQDGLVGRGLVGRGLVGRGLVGRGLVGRGLVGRGLVGRGLVGRGLV 377
QY 383 SQGSELEYMAVYSDITNASLHIGTHLKNLCPRLVLLDHEEKITDPLDNGVRALLR 442
DB 378 AVGCELEYMAVYSDITNASLHIGTHLKNLCPRLVLLDHEEKITDPLDNGVRALLR 437
QY 443 GCDKLRRFALYLRGGLTVDGLGYTGQYSPNVRWMLLVGVGESDAGLLEFAKGCPSLOKL 502
DB 438 NCVKLRREFALYLRGGLTVDGLGYTGQYSPNVRWMLLVGVGESDAGLLEFAKGCPSLOKL 497
QY 503 EMRGCLFESERALAAVQTLNLSRLVWQGVSPSGRDLVWAPFNNIIPG-RKVA 561
DB 498 ELRGCC-FSERALAAVQTLNLSRLVWQGVSPSGRDLVWAPFNNIIPG-RKVA 556
QY 562 TINTNPDETIVVHPAHILAYSLAGQSDFFPTVVPVLPDTAT 602
DB 557 YEVMDAGQPCVTHAQVLAAYSLAGRRPDCQWMLVTLHPAS 597

RESULT 4
US-10-310-154-519
; Sequence 519, Application US/10310154
; Publication No. US2003023670A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; APPLICANT: Chomet, Paul S.
; APPLICANT: Adams, Thomas H
; APPLICANT: Ruff, Thomas G.
; APPLICANT: Agarwal, Ameeta K.
; APPLICANT: Ahrens, Jeffrey E.

; APPLICANT: Ball, James A.
; APPLICANT: Banu, G.
; APPLICANT: Bell, Erin
; APPLICANT: Boddupalli, Raghava
; APPLICANT: Deikman, Jill
; APPLICANT: Deng, Molian
; APPLICANT: Dong, Jinzhao
; APPLICANT: Duff, Stephen M.
; APPLICANT: Galligan, Meghan M.
; APPLICANT: Hinchey, Brenda S.
; APPLICANT: Huang, Shihshieh
; APPLICANT: Johnson, G. Richard
; APPLICANT: Jung, Vincent
; APPLICANT: Kretzmer, Keith A.
; APPLICANT: Laccetti, Lucille B.
; APPLICANT: Lai, Chao-Qiang
; APPLICANT: Lee, Gary
; APPLICANT: Lin, Jie-Yi
; APPLICANT: Liu, Jingdong
; APPLICANT: Lu, Bin
; APPLICANT: Luethy, Michael M.
; APPLICANT: Lund, Adrian
; APPLICANT: Madson, Linda L.
; APPLICANT: Malloy, Kathleen A.
; APPLICANT: McKiel, Christine L.
; APPLICANT: Miller, Philip W.
; APPLICANT: Padmavathi, Manchikanti
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Start, William G.
; APPLICANT: Tennesen, Dan
; APPLICANT: Vidya, K.R.
; APPLICANT: Wang, Haiyuan
; APPLICANT: Xin, Zhaoguo
; APPLICANT: Xu, Nanfei
; APPLICANT: Yang, Chunzhi
; APPLICANT: Zeng, Xiaoping
; APPLICANT: Zhang, Qiang
; APPLICANT: Zhao, Yajuan
; APPLICANT: Zhou, Li
; TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants
; FILE REFERENCE: 38-15(52796)B
; CURRENT APPLICATION NUMBER: US/10/310,154
; CURRENT FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: 60/337,358
; PRIOR FILING DATE: 2001-12-04
; NUMBER OF SEQ ID NOS: 736
; SEQ ID NO 519
; LENGTH: 591
; TYPE: PRT
; ORGANISM: Zea mays
; US-10-310-154-519

Query Match      53.5%; Score 1701.5; DB 15; Length 591;
Best Local Similarity 57.0%; Pred. No. 7.7e-164;
Matches 330; Conservative 95; Mismatches 135; Indels 19; Gaps 7;

QY 28 VDVVLDVIVYIDDPKRDVAVSOVCREWYELDSLTRKHVTIALCYTTTPARLRPPH 87
DB 20 IPDTALGLVWGYVEDPDRDAISLVCRHWCVRDALSRKHVTVMAYSTTPELFRFPCL 79
QY 88 ESKLKGKPRAMFNLIPEDMGHHVTPWVKESIQYFDCLSLHFRMIVKDSLDQNLARD 147
DB 80 ESKLKGKPRAMFNLIPEDMGHHVTPWVKESIQYFDCLSLHFRMIVKDSLDQNLARD 139
QY 148 RGHVHLAKDKCSGFTTDLGPHGRFCKSRVLFLESSILEKDG-EWLHELALNTVL 206
DB 140 RGHVHLAKDKCSGFTTDLGPHGRFCKSRVLFLESSILEKDG-EWLHELALNTVL 199
QY 207 ETNLFYLTIDIAVKIE--DLELLAKNCPNLVSKLTDCEILDVNFPHKASALEEFCGGTY 266
DB 200 ETNLFYLTIDIAVKIE--DLELLAKNCPNLVSKLTDCEILDVNFPHKASALEEFCGGTY 258
QY 267 BEPE-----RYSAISLPAKLCRLGLTYIGKNELPIVFMFAAVLKKDLLEYAMLTDEHOML 321
```

Db 259 EQGQPVASRNENYFFPSLHLSILYMGNTDMQILFFYATALKKLDLQFTLSTEDHCQ 318  
Qy 322 LIQCPNLVLETRNVIGDRGLVGRCKLRIRTERGDDQ-GMEDEGTVSHRGLI 380  
Db 319 IVQCSNLETLEVRVIGDRGLVQVQAQCKLHURVERGDDQDGLDEQGRISQVGLM 378  
Qy 381 ALSQGCSELYMAVVSDDITNASLEHIGTHLKNLCDFRLVLLDHEEKITDPLDNGVRAL 440  
Db 379 AIAQCPELTWAIHVSDITNALEAVGTCSKNLDFRLVLLDREAHITELPLDNGVRAL 438  
Qy 441 LRGCCKLRARFALYLRGGTDLVGLYIGQYSPNVWMLLGVGSDAGLLFAKGCPSLQ 500  
Db 439 LRGCCKLRARFALYLRGGTDLVGLYIGQYSPNVWMLLGVGSDAGLLFAKGCPSLQ 498  
Qy 501 KLEMGCLFFSERALAVATOLTSRLYLWVGQYSPSGRDLVWARPFWNIELPSRKV 560  
Db 499 KLEVRGCL-FSEHALALALQKLSRLYLWVGQYSPSGRDLVWARPFWNIELPSRKV 553  
Qy 561 AINTNPD-ETVVVEHPAHILAYISLAGQSDPDPDTVPVL 598  
Db 554 -----PDQDEPCEHKKRQLLAYISLAGRTDCPPSVTL 587

RESULT 5  
US-10-425-114-49508  
; Sequence 49508, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaaka, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 49508  
; LENGTH: 313  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 701002452\_FLI.pcp  
US-10-425-114-49508

Query Match 49.8%; Score 1584; DB 12; Length 313;  
Best Local Similarity 100.0%; Pred. No. 2.7e-152;  
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 305 KKLDLLYAMLDTHDCHMLIORCPNLEVTNRNIGDRGLVGRCKLRKLRIRTERGDD 364  
Db 12 KKLDLLYAMLDTHDCHMLIORCPNLEVTNRNIGDRGLVGRCKLRKLRIRTERGDD 71  
Qy 365 QGMEDEGTVSHRGLIALSQGCSELYMAVVSDDITNASLEHIGTHLKNLCDFRLVLLDH 424  
Db 72 QGMEDEGTVSHRGLIALSQGCSELYMAVVSDDITNASLEHIGTHLKNLCDFRLVLLDH 131  
Qy 425 EEKTTDPLDNGVRALLRGCDKLRFPALYLRGGTDLVGLYIGQYSPNVWMLLGVGVE 484  
Db 132 EEKTTDPLDNGVRALLRGCDKLRFPALYLRGGTDLVGLYIGQYSPNVWMLLGVGVE 191  
Qy 485 SDAGLLFAKGCPSIQKLEMGCLFFSERALAVATOLTSRLYLWVGQYSPSGRDLV 544  
Db 192 SDAGLLFAKGCPSIQKLEMGCLFFSERALAVATOLTSRLYLWVGQYSPSGRDLV 251  
Qy 545 MARFPWNIELIPSRKVAINTNPDETIVVVEHPAHILAYISLAGQSDPDPDTVPVLDTATCV 604  
Db 252 MARFPWNIELIPSRKVAINTNPDETIVVVEHPAHILAYISLAGQSDPDPDTVPVLDTATCV 311

Qy 505 DT 506  
Db 312 DT 313  
RESULT 6  
US-10-425-114-56666  
; Sequence 56666, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaaka, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 56666  
; LENGTH: 425  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-ZMFLB73276H05\_FLI.pcp  
US-10-425-114-56666

Query Match 41.1%; Score 1305.5; DB 12; Length 425;  
Best Local Similarity 61.0%; Pred. No. 1e-123;  
Matches 261; Conservative 64; Mismatches 92; Indels 11; Gaps 7;  
Qy 181 LFESESSILEKDGWMLHIALNNTVETLNFILTDIAVVKIE--DLELLAKNCPNLVSK 238  
Db 2 LFESESSILEKDGWMLHIALNNTVETLNFILTDIAVVKIE--DLELLAKNCPNLVSK 58  
Qy 239 LTDCEILDVNFKHASALEEFCGGTYNR--SPERYSALSIPAKLCRL-GLTVIGKNELP 295  
Db 59 MGDCLSDSLIGFPQTSKALQEFAGGAFVGGTYTKEKVIFFPRCLCFGLGTFMGKNMP 118  
Qy 296 IVFMFAAVLKLDLLYAMLDTHDCHMLIORCPNLEVTNRNIGDRGLVGRCKLRKLR 355  
Db 119 VIFPSTMLKLDLQFTLTTHDCHMLIAKCPNLVLEVRNIGDRGLVGRCKLRKLR 178  
Qy 356 LRTERGDDQGMEDBEGTVSHRGLIALSQGCSELYMAVVSDDITNASLEHIGTHLKNLC 415  
Db 179 LRTERGDDQGMEDBEGTVSHRGLIALSQGCSELYMAVVSDDITNASLEHIGTHLKNLC 237  
Qy 416 DFLVLLDHEEKITDPLDNGVRALLRGCDKLRFPALYLRGGTDLVGLYIGQYSPNV 475  
Db 238 DFLVLLDHEEKITDPLDNGVRALLRGCDKLRFPALYLRGGTDLVGLYIGQYSPNV 297  
Qy 476 WMLLVGVGSDAGLLFAKGCPSIQKLEMGCLFFSERALAVATOLTSRLYLWVGQY 535  
Db 298 YMLLVGVGSDAGLLFAKGCPSIQKLEMGCLFFSERALAVATOLTSRLYLWVGQY 356  
Qy 536 SPSEGRDLVWARPFWNIELIPS-RKVAINTNPDETIVVVEHPAHILAYISLAGQSDPDP 594  
Db 357 SRGQDLMLMARPYWNIETAPPESAYRVWADQPCVDTHAQVLAAYISLAGRPPCPQW 416  
Qy 595 VVPLDTAT 602  
Db 417 LVTLHPAS 424  
RESULT 7  
US-10-425-114-69709  
; Sequence 69709, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong

```
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 69709
; LENGTH: 410
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17190F07_FLI.pep
; US-10-425-114-69709

Query Match 40.1%; Score 1275.5; DB 12; Length 410;
Best Local Similarity 61.4%; Pred. No. 1.1e-120;
Matches 254; Conservative 62; Mismatches 87; Indels 11; Gaps 7;

QY 195 WLHMLANNVTLEINLYLNDIAVVKIE--DLELLAKNCNLYSVKLTDCIILLDNFFK 252
Db 1 WLHMLAVNNSVLKINFYMTL---KVEPADLELLAKNCKSLSLKMGDCDLSDLIGFFQ 57
QY 253 HASALEFFCGGTNE--EPERYSAISLPKLCRL-GLTYIGKNELPVFMFAVLKKLDL 309
Db 58 TSXALQBFAGGAFVEGVEYTKIEVIPPRLCFGLGUTFMGKNEMPVIFPYSTMLKKLDL 117
QY 310 LYAMLTEDHGMILQRCNLEVLERNVIGRGLVLRGCKRLRLRIERGDDDDQGMED 369
Db 118 QFTLTTHDHCQLIAKCNLSVLEVRNVIGRGLVEVAATCKLRLRIERGDDDDQGE- 176
QY 370 BEGTVSHRGLIALSQGSELEYMAVYSDITNASLEHIGTHLKNLCDFRLVLDHEKIT 429
Db 177 EGGVSGTGLTAVAVGCRELEYIAAYVSDITNGALESIGTFCKNLYDFRLVLDKOKKIT 236
QY 430 DPLDNGVRALLRGCDKRLRFALVLRGGLTDLVGLVGYGSPNVNRMILGYVGESDAGL 489
Db 237 DPLDNGVRALLRNCVKLRFAFLRPGGLSDVGLVGYGSGNIQYMLLNGVESDNLG 296
QY 490 LEFAGKCPESLOKLEMRGCLFFSERALAVAAVQTLTSLRYLWVQYGVSPSGRDLVMPARPF 549
Db 297 IQFATGCTNLKLELRGCC-FSERALAVAVLQMPSLRYINWQYGRASRTGQDMLMARPY 355
QY 550 WNIELI-PSRKVANTNPDETVEVVEHPAHILAYYSLAGORSDFPDVTVPLDTAT 602
Db 356 WNIEFAPPSPESAYRVNMDGQPCVDTHAQVLAAYVSLAGRRPDCQWLVTLHPAS 409

RESULT 8
US-10-425-114-46788
; Sequence 46788, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 46788
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Zea mays

; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 245726
; LENGTH: 572
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_63922C.1.pep
; US-10-424-599-245726

Query Match 23.9%; Score 759.5; DB 12; Length 572;
Best Local Similarity 33.2%; Pred. No. 6.8e-68;
Matches 194; Conservative 95; Mismatches 256; Indels 39; Gaps 12;

QY 30 DWLDCVPIYIDDPKRDVAVSOVCRWYELDSLTKXHVITIALCYTTTPARLRPPHLES 89
Db 7 DEVIEHIFDVVSVSHDRNALSLVCKSWYRIERCTQRVFIGNCYSITPERLIQRPGLKS 66
QY 90 LKLGKPPAAMFNILPEDWGGHVTWPWKEISQYDCLKSLHFRMIVKDSLDONLARDG 149
Db 67 LTLKGPHFADSLVPYDNGGFGVHPWIBALAKNVKGLLEELKRMVSDSELELSRSET 126
QY 150 HVLHAKLDKCGGFTTDCGLFHIGRCKSLRVLFLFEESILEKDGFWLHELANNVTLET 209
```

```
; FEATURE:
; OTHER INFORMATION: Clone ID: 700800510_FLI.pep
; US-10-425-114-46788

Query Match 37.5%; Score 1193; DB 12; Length 394;
Best Local Similarity 61.7%; Pred. No. 2.6e-112;
Matches 234; Conservative 57; Mismatches 84; Indels 4; Gaps 3;

QY 223 DLELLAKNCNLYSVKLTDCIILLDNVNFPHKASALEFFCGGTNEEPE--RYSAISLPK 280
Db 14 DLELLAKSCKSLTSLKIGDCDLSDLIGFFQSATSLEEFAGGTNGOGELTKYGDVFPSP 73
QY 281 LCELGLTYIGKNELPVFMFAVLKKLDLYAMLTEDHGMILQRCNLEVLERNVIGD 340
Db 74 ICSLGLTFMGANEMPIIFFFSALKKLDLQYTTTETHDHCQLIAKCNLSVLAVERNIGD 133
QY 341 RGLVLRGCKRLRLRIERGDDDDQGMEDBEGTVSHRGLIALSQGSELEYMAVYSDIT 400
Db 134 RGLGVVADTCKLQRLRIERGDDDEGGVQBGVQVGLTAIAVGCRELEYIAAVVSDIT 193
QY 401 NASLEHIGTHLKNLCDFRLVLDHEKITDPLDNGVRALLRGCDKRLREALYLRGGLT 460
Db 194 NGALESIGTFCKKLYDFRLVLDREERITDPLDNGVRALLRGCTMLRRFALYLRPGGLS 253
QY 461 DVGLGYIGQYSPNVNRMILGYVGESDAGLEFAGKCPESLOKLEMRGCLFFSERALAVAAT 520
Db 254 DAGLGYIGQYSGNIQYMLLNGVETDDGLISFALGCNLRKLELRSCC-FSERALALAIL 312
QY 521 QLTSRLYRLWVQYGVSPSGRDLVMPARPFWNIELI-PSRKVANTNPDETVEVVEHPAHIL 579
Db 313 SMESLRYVWVQYKASQTRDMLMARPFWNIEFTPPSSQNAGRLEIDGFCVDSHAQIL 372
QY 580 AYISLAGORSDFPDVTVPL 598
Db 373 AYISLAGKRLDCQSVVTL 391

RESULT 9
US-10-424-599-245726
; Sequence 245726, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 245726
; LENGTH: 572
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_63922C.1.pep
; US-10-424-599-245726

Query Match 23.9%; Score 759.5; DB 12; Length 572;
Best Local Similarity 33.2%; Pred. No. 6.8e-68;
Matches 194; Conservative 95; Mismatches 256; Indels 39; Gaps 12;

QY 30 DWLDCVPIYIDDPKRDVAVSOVCRWYELDSLTKXHVITIALCYTTTPARLRPPHLES 89
Db 7 DEVIEHIFDVVSVSHDRNALSLVCKSWYRIERCTQRVFIGNCYSITPERLIQRPGLKS 66
QY 90 LKLGKPPAAMFNILPEDWGGHVTWPWKEISQYDCLKSLHFRMIVKDSLDONLARDG 149
Db 67 LTLKGPHFADSLVPYDNGGFGVHPWIBALAKNVKGLLEELKRMVSDSELELSRSET 126
QY 150 HVLHAKLDKCGGFTTDCGLFHIGRCKSLRVLFLFEESILEKDGFWLHELANNVTLET 209
```

```

Db      127 H-FKSLVLVCEGFSTDGLAALANCRFLRELDLOENEVEDEHKQWLSCTPDNCTSLVSL 185
QY      210 NFYLTDAVVKIEDLELLAKNCPNLVSKLTDCEILD-LVNFFKHASALEEFCGGTYNEE 268
Db      186 NFACLK-GEVSLGALERFVARSPNLKSLKLNRSVPFDALQRMRAQPSLDLGTIGSLVHD 244
QY      269 PERYSALSIPAKLCRL-----GLTYIGKNELPIVEMFAAVLKKLDLLYAM-LDTEHDC 320
Db      245 PESAYIKLNTIUKCKSITSLSGFLVAPICLAAIYPICNLSLMSLAAGIQGSDLV 304
QY      321 MLIORCNLEVLFRNVIGRGLVIGRCCKRLKRLR-----IERGDDDDQGMEDREGTVSH 376
Db      305 KLIHCVKQLRLIMDCIGDKGLGVWATTCOLQELAVFSPVPGD-----PAAVTE 356
QY      377 RGLIALSGCSELEYMAVYSDITNASLEHIGTHLKNLCDFRILVLLDHEKITDL--PLD 434
Db      357 KGLVAISMGCPEKLSLYFCHQMTNAALITVAKNCPNFIIRFLCILDATKPDPTMQPLD 416
QY      435 NGVRAILRGCDKLRFAIYLRGGITDVLGIGYQYSPNVRWMLLGYVGBSDAGLLEFAK 494
Db      417 EGFCAIVQSCRLRLSL-----SGKLTQVFLYIGYAEKLEMUSIAFAGDGDGMGLYLN 473
QY      495 GCPSLQKLEMRGCLFFSERALAAVAATQLTSLRYLWQYGYVSPSGRDLVLMARPFWMNIEL 554
Db      474 GCKLRKLEIRDCP-FGNMALLTDVGKYETMRSLWSSCEVTGACKLLAKKMPRLNVEI 532
QY      555 IPSRKVATNTPDETVEVVEHPAHIT---LAYISLAGQSDPFDTV 595
Db      533 F-----NENEQEDCSLEDGQVKYKMYLRYTLAGKCKDAPEYV 569

```

## RESULT 10

```

US-10-424-599-234595
; Sequence 234595, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 234595
; LENGTH: 587
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(587)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_53867C.1.pap
US-10-424-599-234595

```

```

Query Match      23.2%; Score 737; DB 12; Length 587;
Best Local Similarity 32.7%; Pred. No. 1.4e-65;
Matches 188; Conservative 108; Mismatches 253; Indels 26; Gaps 14;

QY      33 LDCVIPYIDDPKRDVAVQVCRRWYELDSLTRKHVTIALCYTTTPARLRPPHLESKL 92
Db      15 LEVTFPMNERDRNAISLVCKSWYIERWCRSKVFVGNVCYAVSLMVVKRPPEVTSIAL 74
QY      93 KGKPRAMFNLIPEDEGGHVTVPWKEISQVDFCLKSLHFRFIRMIVKDSLQNLARDRGHVL 152
Db      75 KGKPFADFDLVPGWGCYICPWDAVRSFPCFEQFLKTKWTWITDESLELAKSPKN-F 133
QY      153 HALKDKCSGFTDGLFHIGRCFKSLRVLFLBESSILEKDGEMHELANNVTLETIN 210

```

```

Db      134 KVLVLTSCEGKTPYGLAAIAANCRNLRELDLQSEV-EEDLCGHWSLGHFPDSTYSLVSLN 192
QY      211 FYLTDAVVKIEDLELLAKNCPNLVSKLTDCEILD-LVNFFKHASALEEFCGGTYNEE- 268
Db      193 ISCLN-NEVSLSALERLLGRCPNLRTLRANRAVPLDRPLNLLRAPQLVELGTGVYSTEM 251
QY      269 -PERYSALSIPAKLCR-----LGLTYIGKNELPIVEMFAAVLKKLDLLYAMLTDEHCLM 322
Db      252 RPEVFNLEAAFAFGCKQLKSLSGFMDVLPSTLPAVVPICSRSLTSLNLSYAIIOSSDLIKL 311
QY      323 IQCPNLLEVLTRNVIGRGLVIGRCCKRLKRLRTERGGDDOGMEDEEGTVSHRGLIAL 382
Db      312 ISQPNLLRLWLDYIEDAGLYALAASKDLRELVR-FPSDPPGLEPNV-ALTEQGGLVSV 369
QY      383 SQCSELEYMAVYSDITNASLEHIGTHLKNLCDFRILVLLDHE--EKITDPLDNGVRL 440
Db      370 SEGCPKLSVLYFCRQMSNAALHTIARNRTNLTRFLCIIEPTPTDVLTHEPLDSGFGAI 429
QY      441 LRGCCKLRPALYLRGGITDVLGIGYQYSPNVRWMLLGYVGBSDAGLLEFAKGCPSLQ 500
Db      430 VEQCKLORLSL---SGLLTDRVFEYIGTCGKKLEMLSVAFAGSDGLGHHVLSGCDNLR 486
QY      501 KLEMRGCLFFSERALAAVAATQLTSLRYLWQYGYVSPSGRDLVLMARPFWMNIELIPSKV 560
Db      487 KLEIRDCP-FGDKALLANAEEKLETMRSLWSSCEVSYGACKLLGQKMPRLNVEVIDERG- 544
QY      561 ATNPNDETVEVVEHPAHILAYISLAGQSDPFDTV 595
Db      545 PDSRPDDCPVEK---LYIYRTIAGPRLDMPGFV 575

```

## RESULT 11

```

US-10-424-599-212007
; Sequence 212007, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 212007
; LENGTH: 578
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_33469C.1.pap
US-10-424-599-212007

```

```

Query Match      23.1%; Score 735; DB 12; Length 578;
Best Local Similarity 32.0%; Pred. No. 2.2e-65;
Matches 187; Conservative 107; Mismatches 251; Indels 40; Gaps 13;

QY      30 DVVLDCVIPYIDDPKRDVAVQVCRRWYELDSLTRKHVTIALCYTTTPARLRPPHLES 89
Db      18 DEVLERILLGMLSKRDKSTVSLVCKEWFENAEWRSSRVFIGNCYSVSPILLTRFPFNIRS 77
QY      90 LKLCKPRAMFNLIPEDEGGHVTVPWKEISQVDFCLKSLHFRFIRMIVKDSLQNLARDRG 149
Db      78 VTLGKPRFSDNLPVPAWNGADIHSMVFAEAKPWEELRKRMTVTDESLEFLAKRFP 137
QY      150 HVLHALKDKCSGFTDGLFHIGRCFKSLRVLFLBESSILEKDGEMHELANNVTLETI 209
Db      138 N-FKALLSLSCDGFSTDGLASATCKNLTELDIOENGIEDKSGNLSCFPDSTFSLVL 196
QY      210 NF--YLTDAVVKIEDLELLAKNCPNLVSKLTDCEILD-LVNFFKHASALEEFCGGTYN 266
Db      197 NFANLHND---VNFDALEKLVSRCKSLATLKNVSKVTLEQLQRLIVHVHPQLGELGTGSFS 253

```

[illegible]

```

RESULT 12
US-10-424-599-143159
; Sequence 143159, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 143159
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_100286C.1.pep
US-10-424-599-143159

```

[illegible]

RESULT 13  
US-10-425-114-68189  
; Sequence 68199, Application US/10425114  
; Publication No. US2004003488A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua

```

; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 68189
; LENGTH: 597
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4570-007-All_FLI pep
US-10-425-114-68189

```

Query Match	21.7%	Score 690.5	DB 12	Length 597
Best Local Similarity	31.2%	Fred. No. 7.9e-61		
Matches	181	Conservative 114	Mismatches 251	Indels 35
Gaps	14			
Qy	32	VLDGVPIYDDKDRDAVSOVCRWYELDSLTKHVHTIALCYTTTPARLRFRFPHLSLK	91	
Db	30	VWEHFSFLSHSDRNTVSVLCKWYWEVERLSRAVFGVNCYAVRPERVVLRFENIKALT	89	
Qy	92	LKGKPRAAENFLIPEDWGGCHVTWPVKEISQVDFCLKSLHFRMIKSDSDQLNARD--RG	149	
Db	90	VKGKPHDFNLVPPDWGGYAGPWIDAARSCVGLEELRMKRMVVFENLELLARSELRF	149	
Qy	150	HVLHALKDKCSGFTTDLGPHIGFCKSLAVLFEBSLSLEKDGWEHLELANNTVLETL	209	
Db	150	KV----LVLSICEGFTDGLAAISHCKLLRELDQENDVEDRGPRWL--SPFDSCTSLVSL	205	
Qy	210	NFYLTDTIAVKIE----DLELLAKNCNLVSVKLTDCIELD--LYNFFKHASALEEFCOGT	264	
Db	206	NF-----ACTIGEVNSGALERBELVARSFNLSRLNRSVSDTLLSKILLRAPNLEDLGTGN	260	
Qy	265	YNEB--PERYSAISLPKRLGLTYTGKNE-----LPIVFMFAVLKDLDLLYA-MLDT	316	
Db	261	LTDFQAESYRSLTSALEKCKRLSLSGFWDASPICVPIYPLCHQTLGNLSYTPTLDY	320	
Qy	317	EDHMLQRCNLBEVLETRNVIGDRGLEVLRCGRCKRLKRLIERGDDDDQGMDEBEGTVSH	376	
Db	321	SDLAKMVRVCVKLQRLWLVDCISDKGLQVASSCKDLQELRV---PFSEFVVPGASAVTE	377	
Qy	377	RLTALSQGCSELEYMAVYVSDITNASLEHGTGLKNLCPRLVLDHE--EKITDLPDL	434	
Db	378	EGLVAISGCPKLTSLLYFCHQMTNEALITVANNCPNFIREFLCILEPFKPKEDAMTQGLD	437	
Qy	435	NGVTALLRGCDLRRFPALYLRGGTLTVGLGYGOYSPNVRWMLLGYVGSDDAGLLEPAK	494	
Db	438	EGFGAIVRECKGLRRLSI---SGULTDKVFMVYIGKHAKYLEMLSIGAFAGSDSKGMVDMN	494	
Qy	495	GCPSIQLEMRGCLFFSERALAVAATQUTSLRYLVWQGYGYSPPGRDLLVMVAPFPWNIL	554	
Db	495	GCKNLKLEIRDSP--FGDVALLGNVAKYETMRSLWMSSCNVTLKGCQVLASKMPMLNVEI	553	
Qy	555	IPSRKVANTNPDTVTVVVEHPAHILAYYSLAGQSDPDDTV	595	
Db	554	M--NELDGSSMENHGDLSSKVDKLYVYRTTAGARDDAFNFV	592	

RESULT 14  
US-10-310-154-655  
; Sequence 655, Application US/10310154  
; Publication No. US2003023670A1  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D  
; APPLICANT: Chomet, Paul S  
; APPLICANT: Adams, Thomas H  
; APPLICANT: Ruff, Thomas G  
; APPLICANT: Agarwal, Ameeta K.

Search completed: April 20, 2004, 01:49:29  
Job time : 364 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 22:59:35 ; Search time 19 Seconds  
(without alignments)  
1660.764 Million cell updates/sec

Title: US-10-009-791-22  
Perfect score: 3180  
Sequence: 1 TKTSAFFLTLSRSNWTEE.....QRSDFPDVVPLDTCVDT 606

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	183.5	5.8	461	1 YKK7 CAEEL	P34284 caenorhabdi
2	179	5.6	436	1 FL2L HUMAN	Q96192 homo sapien
3	175	5.5	423	1 FL2L MOUSE	Q8bhl6 mus musculus
4	175	5.5	436	1 FL2L MOUSE	Q9czv8 mus musculus
5	172	5.4	491	1 FL17 HUMAN	Q9ujt9 homo sapien
6	170.5	5.4	423	1 FL2L HUMAN	Q9ukc9 homo sapien
7	154	4.8	463	1 POF2 SCHPO	O74783 schizosacch
8	147.5	4.6	276	1 FL2L RAT	Q9qzh7 rattus norv
9	141	4.4	539	1 FL16 HUMAN	Q8n531 homo sapien
10	139	4.4	630	1 ESA8 TRYEQ	P26337 trypanosoma
11	133	4.2	630	1 ESA8 TRYEB	P23799 trypanosoma
12	132	4.2	565	1 RAD7 YEAST	P06779 saccharomyc
13	131.5	4.1	1061	1 NA12 HUMAN	P39046 homo sapien
14	127.5	4.0	1151	1 GRR1 YEAST	P24814 saccharomyc
15	126.5	4.0	621	1 FL16 HUMAN	Q9uka2 homo sapien
16	120.5	3.8	535	1 FL16 MOUSE	Q9qzw0 mus musculus
17	119	3.7	1024	1 POPC BALSO	Q9rbs2 raltstonia s
18	118	3.7	699	1 YQOA CAEEL	Q09299 caenorhabdi
19	117.5	3.7	1093	1 NA14 HUMAN	Q86w24 homo sapien
20	115	3.6	434	1 FX3B HUMAN	Q9ukt6 homo sapien
21	114.5	3.6	434	1 FX3B MOUSE	Q8bfz4 mus musculus
22	110.5	3.5	326	1 FX1C HUMAN	Q9nuk8 homo sapien
23	110.5	3.5	424	1 SKP2 HUMAN	Q13309 homo sapien
24	109	3.4	326	1 FX1C MOUSE	Q9epx5 mus musculus
25	108	3.4	424	1 SKP2 MOUSE	Q92023 mus musculus
26	106.5	3.3	880	1 GYRA HAEIN	P43700 haemophilus
27	105.5	3.3	223	1 FX19 HUMAN	Q8wv35 homo sapien
28	105.5	3.3	1034	1 CLS1 HUMAN	Q96p20 homo sapien
29	104.5	3.3	980	1 CLV1 ARATH	O9syq8 arabidopsis
30	103	3.2	859	1 MUTS AQUAE	O66552 aquifex aeo
31	103	3.2	868	1 MUTS XYLFT	Q87c18 xyellia fas
32	103	3.2	1039	1 YR71 CAEEL	Q09564 caenorhabdi
33	102.5	3.2	566	1 Y397 MYCGE	P47637 mycoplasma

34	102.5	3.2	1024	1 CARC HUMAN	Q9npp4 homo sapien
35	102.5	3.2	1131	1 PHVA SOYEN	P42500 glycine max
36	102	3.2	582	1 SHO2 MOUSE	O88520 mus musculus
37	102	3.2	834	1 AIIM YEAST	P03875 saccharomyc
38	101.5	3.2	456	1 RINI PIG	P10775 sus scrofa
39	101	3.2	460	1 RINI HUMAN	P13489 homo sapien
40	101	3.2	681	1 COG2 CAEEL	Q21444 caenorhabdi
41	101	3.2	951	1 LGR4 HUMAN	Q9bxb1 homo sapien
42	100.5	3.2	2052	1 UBRB SCHPO	O13731 schizosacch
43	100	3.1	446	1 AK RICPR	Q9zci7 rickettsia
44	100	3.1	2364	1 SPQO HUMAN	Q01082 homo sapien
45	99	3.1	399	1 R24L ARATH	Q9c646 arabidopsis

ALIGNMENTS

RESULT 1  
YKK7 CAEEL STANDARD; PRT; 461 AA.  
AC P34284;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Hypothetical F-box/LRR-repeat protein C02F5.7 in chromosome III.  
GN C02F5.7.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae; Rhabditidae;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RS SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RX MEDLINE=94150718; PubMed=7906398;  
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
RA Craxton L., Dear S., Du Z., Durbin R., Favello A., Fraser A.,  
RA Fulton M., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,  
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,  
RA Lathille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,  
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,  
RA Sims M., Smalton N., Smith A., Smith M., Sonhammer E., Staden R.,  
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,  
RA Waterston R., Watson A., Weinstein L., Wilkinson-Sproat J.,  
RA Wohldman P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
RL Nature 368:32-38(1994).  
[2]

Waterston R.;  
Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

CC - SIMILARITY: Contains 1 F-box domain.

CC - SIMILARITY: Contains 11 leucine-rich (LRR) repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; L14745; AAA27922.2; -;  
DR PIR; S44609; S44609;  
DR WormPep; C02F5.7; CE24780.

DR InterPro; IPR001810; F-box.

DR InterPro; IPR001611; LRR.

DR InterPro; IPR007089; LRR\_cyb.

DR InterPro; IPR006553; LRR\_cyb\_sub.

DR Pfam; PF00646; F-box; 1.

DR Pfam; PF00560; LRR; 2.

DR SMART; SM00256; FBOX; 1.

DR SMART; SM00367; LRR CC; 5.  
DR PROSITE; PS50181; FBOX; 1.  
KW Hypothetical protein; Ubl conjugation pathway; Repeat;  
KW Leucine-rich repeat.  
FT REPEAT 54 100 F-BOX.  
LRR 1. 122 147  
LRR 2. 148 173  
LRR 3. 174 199  
LRR 4. 200 225  
LRR 5. 226 251  
LRR 6. 252 277  
LRR 7. 278 303  
LRR 8. 304 329  
LRR 9. 330 355  
LRR 10. 356 381  
LRR 11. 408 433  
SQ SEQUENCE 461 AA; BB28C98A5CA7B1D5 CRC64;  
Query Match 5.8%; Score 183.5; DB 1; Length 461;  
Best Local Similarity 22.0%; Pred. NO. 1.2e-06;  
Matches 98; Conservative 63; Mismatches 134; Indels 151; Gaps 17;  
QY 50 SQVCRVVEL-----DSLTR-----KHVTIALCVTTTPAR 79  
Db 81 AQVCRSIIALDGSNWQRDLFTFQDVKTAUVENLARRCGGLKSLKGCENVHDSA 140  
QY 80 LR---RRPHLESILKLGKPRAMFNLIPEDWGHVTFWVKETISQYFDCLKSLHFRMIV 136  
Db 141 LRTFTSRCPNLEHLSL-----YRCKRV 162  
QY 137 KSDLQNLARQGHVILHALKDKSCGFTTDLGFLHGRCKSLRVLFLEESSILEKDGWL 196  
Db 163 TDASCENLGR-YCHKLYNINLENCSSITDRAMKYIGDGCPLNLSINWCDAIQDGV-- 219  
QY 197 HELALNN-TVLETNFY---LTDAIVVKIDLELLAKNCPNLVSVKLTDCIEIDLNVFF 251  
Db 220 -QILSNCKSLDTILRCCEGLTENFVGSVAHGAIKNLNLCQFQITDITVQNIAN-- 276  
QY 252 KHASALEFCGGTYNEEPERYSAISLPKLCRLGLTYIGKNELPIVFMFAVKKLLDLY 311  
Db 277 -GATALEYLCMNCNQISDR-----SLVSLG----- 301  
QY 312 AMLOTEDEHMLIQRCPNLEVLETR--NVIGDGLVLCRCCKRLKRLRIERGDDQGMED 369  
Db 302 -----QSHNLIKLVLELSCTLIUGNGFPIPLARGCRQLERLDME----- 339  
QY 370 EEGTVSHRGLIALSGCSELEYMAV-YVSDITNASLEHIGT-HLKNLDCFRVLVLDHEEK 427  
Db 340 DCSLSDHTINSIANNCTALRELSLSHCELITDESIGNLASKHRETL--NVLELDNCPQ 396  
QY 428 ITDPLDNGVALLRGCDKLRFPALY 453  
Db 397 LTDSTLSH-----LRHCKALKRIDLY 417  
RESULT 2  
FL2L HUMAN STANDARD; PRT; 436 AA.  
AC Q961G2;  
DT 15-MAR-2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DE 15-MAR-2004 (Rel. 43, Last annotation update)  
DE F-box/LRR-repeat protein 2-like.  
GN FBL2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Colon;  
RX MEDLINE=22398257; PubMed=12477932;  
RA Strausberg R.L.; Feingold E.A.; Grouse L.H.; Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
Diachenko L., Marusina K., Farmer A.F., Rubin G.M., Hong L.,  
Stapleton M., Soares M.B., Bonaldo M.F., Carninci P., Prange C.,  
Brownstein M.J., Usdin T.B., Toshiyuki S., Abramson R.D., Mullany S.J.,  
Raha S.S., Loquellano N.A., Peters G.J., Malek J.A., Gunaratne P.H.,  
Bosak S.A., McEwan P.J., McKernan K.J., Garcia A.M., Gay L.J., Hulyk S.W.,  
Richards S., Worley K.C., Hale S., Sodergran E.J., Lu X., Gibbs R.A.,  
Villalon D.K., Muzny D.M., Sodergran E.J., Lu X., Gibbs R.A.,  
Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
Butterfield Y.S.N., Krzywinski M.I., Skalka U., Smalilus D.E.,  
Scherer A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -I- FUNCTION: Probably recognizes and binds to some phosphorylated  
proteins and promotes their ubiquitination and degradation.  
CC -I- SUBUNIT: Interacts with SKP1. Part of a SCF (SKP1-cullin-F-box)  
CC -I- SUBUNIT: Interacts with SKP1. Part of a SCF (SKP1-cullin-F-box)  
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -I- SIMILARITY: Contains 1 F-box domain.  
CC -I- SIMILARITY: Contains 12 leucine-rich (LRR) repeats.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
CC EMBL; BC007557; AA07557.2; -;  
DR InterPro; IPR001810; F-box.  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR007089; LRR\_CYS.  
DR InterPro; IPR006553; LRR\_CYS\_sub.  
DR Pfam; PF00646; F-box; 1.  
DR Pfam; PF00560; LRR; 2.  
DR SMART; SM00256; FBOX; 1.  
DR SMART; SM00367; LRR CC; 12.  
DR PROSITE; PS50181; FBOX; 1.  
KW Ubl conjugation pathway; Leucine-rich repeat; Repeat.  
FT DOMAIN 22 68 F-BOX.  
REPEAT 90 115 LRR 1.  
REPEAT 116 141 LRR 2.  
REPEAT 142 167 LRR 3.  
REPEAT 168 193 LRR 4.  
REPEAT 194 219 LRR 5.  
REPEAT 220 245 LRR 6.  
REPEAT 246 271 LRR 7.  
REPEAT 272 297 LRR 8.  
REPEAT 298 323 LRR 9.  
REPEAT 324 349 LRR 10.  
REPEAT 353 377 LRR 11.  
REPEAT 378 403 LRR 12.  
SQ SEQUENCE 436 AA; 48423 MW; 39CD04A505C8CE3E CRC64;  
Query Match 5.6%; Score 179; DB 1; Length 436;  
Best Local Similarity 22.5%; Pred. NO. 2.5e-06;  
Matches 74; Conservative 47; Mismatches 114; Indels 94; Gaps 11;  
QY 151 VLHALKDKSCGFTTDLGFLHGRCKSLRVLFLEESSILEKDGWLHELALNNTVLETLN 210  
Db 170 LLEQLNISWCDDQVTKDGIQALVRGGGKALFLKCTQLD----- 210  
QY 211 FYLTDAIVVKIEDLELLAKNCPNLVSVKLTDC-EILDVNFHFKHASALEEFCGGTYNEEP 269  
Db 211 -----EALKYIGNHCPVLTNLNLTCLQITD-----EGLITTCRGCHKLQ- 250



245	271	LRR 7.
246	272	LRR 8.
247	273	LRR 9.
248	274	LRR 10.
249	275	LRR 11.
250	276	LRR 12.
251	277	Missing (in isoform 2).
252	278	/FTID=VS9P008968.
253	279	L -> P (IN REF. 1; BAB28039).
254	280	SEQUENCE 436 AA; 48390 MW; C800E1861AF21BC3 CRC64;
255	281	
256	282	
257	283	
258	284	
259	285	
260	286	
261	287	
262	288	
263	289	
264	290	
265	291	
266	292	
267	293	
268	294	
269	295	
270	296	
271	297	
272	298	
273	299	
274	300	
275	301	
276	302	
277	303	
278	304	
279	305	
280	306	
281	307	
282	308	
283	309	
284	310	
285	311	
286	312	
287	313	
288	314	
289	315	
290	316	
291	317	
292	318	
293	319	
294	320	
295	321	
296	322	
297	323	
298	324	
299	325	
300	326	
301	327	
302	328	
303	329	
304	330	
305	331	
306	332	
307	333	
308	334	
309	335	
310	336	
311	337	
312	338	
313	339	
314	340	
315	341	
316	342	
317	343	
318	344	
319	345	
320	346	
321	347	
322	348	
323	349	
324	350	
325	351	
326	352	
327	353	
328	354	
329	355	
330	356	
331	357	
332	358	
333	359	
334	360	
335	361	
336	362	
337	363	
338	364	
339	365	
340	366	
341	367	
342	368	
343	369	
344	370	
345	371	
346	372	
347	373	
348	374	
349	375	
350	376	
351	377	
352	378	
353	379	
354	380	
355	381	
356	382	
357	383	
358	384	
359	385	
360	386	
361	387	
362	388	
363	389	
364	390	
365	391	
366	392	
367	393	
368	394	
369	395	
370	396	
371	397	
372	398	
373	399	
374	400	
375	401	
376	402	
377	403	
378	404	
379	405	
380	406	
381	407	
382	408	
383	409	
384	410	
385	411	
386	412	
387	413	
388	414	
389	415	
390	416	
391	417	
392	418	
393	419	
394	420	
395	421	
396	422	
397	423	
398		

QY	402	ASLEHGTGTHKNL	CD	FL	VL	LD	DE	HE	KI	T	D	I	P	L	D	N	G	V	R	A	L	L	R	G	C	D	K	R	R	F	A	L	R	-	R	G	-	G	450												
Db	343	REIAKLESRL	-----	RYLSIAH	CG	R	V	T	D	V	-----	GIRV	V	A	K	Y	C	S	K	L	R	-----	Y	N	A	R	G	C	E	G	348																				
QY	459	LTDVGLG	Y	IG	O	Y	S	P	N	V	R	M	L	L	G	V	Y	-	E	S	D	A	G	L	L	E	F	A	K	G	C	P	S	I	O	K	L	E	M	R	G	C	L	F	S	E	R	A	L	V	517
Db	389	ITDHGVEYL	A	K	N	C	T	K	L	S	D	I	G	K	P	L	V	S	D	T	G	L	E	C	L	A	N	C	F	N	L	K	R	L	S	A	C	S	E	I	T	G	O	G	L	Q	I	448			
QY	518	AATQLTSL	R	Y	L	W	Q	Y	G	V	S	P	S	R	D	L	L	V	N	A	R	F	F	N	I	E	L	I	P	S	R	K	V	A	T	N	P	566													
Db	449	VAA	N	C	F	D	L	Q	T	L	N	V	D	C	E	V	S	-----	V	E	A	L	R	F	-	V	K	R	C	K	R	V	I	E	H	T	N	P	488												

RESULT 6

FL2 HUMAN

FL2 ID

FL2 HUMAN

STANDARD

PRT; 423 AA.

AC

Q9UKQ9; Q9UKQ8; Q9UKA5; Q9Y3Y9;

DT

15-MAR-2004 (Rel. 43, Created)

DT

15-MAR-2004 (Rel. 43, Last sequence update)

DT

15-MAR-2004 (Rel. 43, Last annotation update)

DE

F-box/LRR-repeat protein 2 (F-box and leucine-rich repeat protein 2)

DE

(F-box protein FBL2) (F-box protein FBL3).

OS

FBL2 OR FBL3

GN

Homo sapiens (Human)

OC

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX

NCBI\_TaxId=9606;

[1]

SEQUENCE FROM N.A.

MDLINE=2003060; PubMed=10531035;

RA

Cenciarelli C., Chiaur D.S., Guardavaccaro D., Parks W., Vidal M.,

RA

Pagano M.;

RA

"Identification of a family of human F-box proteins.";

RL

Curr. Biol. 9:1177-1179(1999).

[2]

SEQUENCE FROM N.A., INTERACTION WITH SKPL, AND TISSUE SPECIFICITY.

MDLINE=2003061; PubMed=10531037;

RA

Winston J.T., Kepp D.M., Zhu C., Elledge S.J., Harper J.W.;

RA

"A family of mammalian F-box proteins.";

RL

Curr. Biol. 9:1180-1182(1999).

[3]

SEQUENCE FROM N.A.

MDLINE=20399565; PubMed=10945468;

RA

Ilyin G.P., Riialand M., Pigeon C., Gugen-Guillouzo C.;

RA

"cDNA cloning and expression analysis of new members of the mammalian

RT

F-box protein family.";

RL

Genomics 67:40-47(2000).

[4]

SEQUENCE FROM N.A.

RA

Isegai T., Oka T., Hayaashi K., Sugiyama T., Otsuki T., Suzuki Y.,

RA

Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,

RA

Tanase T., Nomura Y., Toshiya S., Komai F., Hara R., Takeuchi K.,

RA

Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,

RA

Wakamatsu A., Nakamura Y., Nagahara K., Masuho Y., Oshima A.;

RT

"NDO human cDNA sequencing project.";

RL

Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

[5]

SEQUENCE FROM N.A.

RA

TESSUE=Brain;

RC

MDLINE=22388257; PubMed=12477932;

RA

Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA

Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA

Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA

Hopkins R.F., Jordan K., Moore T., Max S.I., Wang J., Hsieh F.,

RA

Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong E.,

RA

Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Sheetz T.E.,

RA

Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA

Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA

Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA

Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA

Vallalon D., Muzny D.N., Sodergren E.J., Lu X., Gibbs R.A.,

RA

Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grilwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzyzinski M.I., Skalska U., Smalhus D.E.,  
 RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903 (2002).  
 RN [6]  
 RP SEQUENCE OF 174-423 FROM N.A.  
 RC TISSUE=Brain;  
 RA Wambutt R., Heubner D., Mewes H.-W., Gassenhuber J., Wiemann S.;  
 RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Probably recognizes and binds to some phosphorylated  
 CC proteins and promotes their ubiquitination and degradation.  
 CC -!- SUBUNIT: Interacts with SKP1. Part of a SCF (SKP1-cullin-F-box)  
 CC protein ligase complex (By similarity).  
 CC -!- TISSUE SPECIFICITY: Expressed in brain, heart, kidney, liver,  
 CC lung, pancreas and placenta.  
 CC -!- SIMILARITY: Contains 1 F-box domain.  
 CC -!- SIMILARITY: Contains 11 leucine-rich (LRR) repeats.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC  
 CC -----  
 CC EMBL; AF174589; AAF04510.1; -  
 CC EMBL; AF174589; AAF04510.1; ALT\_INIT.  
 CC EMBL; AF186273; AAD56248.1; -  
 CC EMBL; AF186273; AAD56248.1; -  
 CC EMBL; AK001438; BAA91691.1; -  
 CC EMBL; BC031556; AAH31556.1; -  
 CC EMBL; AL049953; CAB43222.1; -  
 CC EMBL; HGNC:13598; FBX12.  
 CC MIM; 605652; -  
 CC InterPro; IPR001810; F-box.  
 CC InterPro; IPR001611; LRR.  
 CC InterPro; IPR007089; LRR\_cys.  
 CC InterPro; IPR006553; LRR\_cys\_sub.  
 CC Pfam; PF00646; F-box; 1.  
 CC Pfam; PF00560; LRR; 3.  
 CC SMART; SM00256; FBOX; 1.  
 CC SMART; SM00367; LRR\_CC; 2.  
 CC PROSITE; PS50181; FBOX; 1.  
 CC KW Ub conjugation pathway; Leucine-rich repeat; Repeat.  
 FT DOMAIN 9 55 F-BOX.  
 FT REPEAT 77 102 LRR 1.  
 FT REPEAT 103 128 LRR 2.  
 FT REPEAT 129 154 LRR 3.  
 FT REPEAT 155 180 LRR 4.  
 FT REPEAT 181 206 LRR 5.  
 FT REPEAT 207 232 LRR 6.  
 FT REPEAT 233 258 LRR 7.  
 FT REPEAT 259 284 LRR 8.  
 FT REPEAT 285 310 LRR 9.  
 FT REPEAT 311 336 LRR 10.  
 FT REPEAT 365 391 LRR 11.  
 FT CONFLICT 62 62 I -> T (IN REF. 4).  
 FT CONFLICT 77 77 G -> V (IN REF. 1).  
 FT CONFLICT 81 81 R -> K (IN REF. 4).  
 FT CONFLICT 174 174 I -> W (IN REF. 5).  
 FT CONFLICT 320 320 S -> P (IN REF. 6).  
 SQ SEQUENCE 423 AA; 47073 MW; BEGF824385121B4E CRC64;  
 Query Match 5.4%; Score 170.5; DB 1; Length 423;  
 Best Local Similarity 23.4%; Pred. No. 1.1e-05;  
 Matches 85; Conservative 54; Mismatches 136; Indels 89; Gaps 15;  
 126 LKSLHFRMI-VKSDSLQNLARDGHVLHKLKDCSGFTTDGLFHGRFCKSLRVLFLF 184

Db 80 LRKLSRGICIGVGDSSIKTFAQNCRIEH-LNLANGCTKTIIDTSCYSLSRFCSKXKLHLDLT 138  
 QY 185 ES-SILEKGEWHEALANNVTLETINFYITDIADVVKIEDLELLAKNCPNLVSVKLTDC- 242  
 Db 139 SCVSIITNSLSKIGISEGCRN--LEYLNLSWCD--QITKDGIEALVRCRGKALLRGCT 193  
 QY 243 ----EILDVNFPKHA-----SALLEFCGGTVNEEPERYSAISLP--AKLC 282  
 Db 194 QLEDEALKHIONQCHELVSLNLSQCSRITDEGVVQICRGCH-----RLQALCLSGCSNLT 248  
 QY 283 RLGLTITGKNEPLFVFWFAAVLKKLLDLYAMLTDEHMLIORCPNLEVLTRNV--IGD 340  
 Db 249 DASITALGLN-----CPRLQILEAARCSHLTD 275  
 QY 341 RLGLVGRCCRLKRLRIERIGDDDDQMEDEGVSHRGLIALSGCSELEYMAV-YVSDI 399  
 Db 276 AGFTLLARNCHELEKMDLE-----ECILITDSTLIQLSHCPKQLQALSUSHCCLI 325  
 QY 400 TNASLEHIGTHLKNLDCDFRLVLDHBEKITDLPDNGVRALLRGCDKLRFFALY----LR 455  
 Db 326 TDGILHLSNSTGCHERLRLVLELDNCLLITDVALEH-----LENCRGLERLDYDCQV 380  
 QY 456 RGGL 459  
 Db 381 RAGI 384

RESULT 7  
 ID\_POF2 SCHPO STANDARD; PRT; 463 AA.  
 AC 074783;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DE 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE F-box/LRR-repeat protein 2 (F-box and leucine-rich repeat protein 2)  
 GN (F-box protein pof2).  
 GN POF2 OR SPBC25B2.11.  
 OS Schizosaccharomyces pombe (Fission Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RA MEDLINE=21669955; PubMed=11809834;  
 RA Katayama S., Kitamura K., Lehmann A., Nikaido O., Toda T.;  
 RT "Fission Yeast F-box protein Pof3 is required for genome integrity and  
 RT telomere function.";  
 RL Mol. Biol. Cell 13:211-224 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RA MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jørgensen K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeh C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Voickart G., Aert R., Robben J., Grymonprez B.,  
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Borzym K., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Eger P., Zimmermann W., Wedler H., Reinhardt R., Pohl T.M.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Laureau V., Motier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode J.,  
 RA Dega R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito G.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerretti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Useery D., Barrell B.G., Nurse P.,  
 RT "The genome sequence of *Schizosaccharomyces pombe*."  
 RL Nature 415:871-880(2002).  
 CC -!- SIMILARITY: Contains 1 F-box domain.  
 CC -!- SIMILARITY: Contains 8 leucine-rich (LRR) repeats.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; AB062778; BAB50689.1; -  
 DR EMBL; AL031853; CAA21269.1; -  
 DR PIR; T39987; T39987.  
 DR GeneDB SPombe; SPBC25B2.11; -  
 DR InterPro; IPR001810; F-box.  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR007089; LRR cys.  
 DR InterPro; IPR006553; LRR cys sub.  
 DR InterPro; IPR008945; SKP1\_Skp2.  
 DR Pfam; PF00560; LRR; 1.  
 DR SMART; SM00367; LRR\_CC; 1.  
 DR PROSITE; PS50181; FBOX; 1.  
 KW Ub1 conjugation pathway; Repeat; Leucine-rich repeat.  
 FT DOMAIN 1 42 F-BOX.  
 FT REPEAT 93 118 LRR 1.  
 FT REPEAT 145 170 LRR 2.  
 FT REPEAT 171 196 LRR 3.  
 FT REPEAT 276 301 LRR 4.  
 FT REPEAT 302 327 LRR 5.  
 FT REPEAT 328 353 LRR 6.  
 FT REPEAT 354 378 LRR 7.  
 FT REPEAT 380 405 LRR 8.  
 SQ SEQUENCE 463 AA; 52360 MW; 07435B3C2BBB2073 CRC64;  
 Query Match 4.8%; Score 154; DB 1; Length 463;  
 Best Local Similarity 20.5%; Pred. NO. 0.00024;  
 Matches 97; Conservative 70; Mismatches 145; Indels 162; Gaps 22;  
 QY 86 HLESILKLGKPRAMFNILIPEDWGHHVTP--VWK-----EISQYFDCLK-----SLHF 131  
 DB 14 VLEADELRCK-----STVCTSRNFIPTLWEKVVFQNEAQLNFFDTLQYSKDVSYF 67  
 QY 132 RMIVKDSLDQLNARDGHVHLAKDKSGFTTGLFHIGRFCKSLRVLFLEESSILEK 191  
 DB 68 R-----YLFKLCNSRVKFLTKD--HLMLMTLATGISRLNLSGCTRI 107  
 QY 192 DGEWLHELALNVTLETINLYTDAVVKIELELLAKNCPLNVLKLTDCBILDVNVFF 251  
 DB 108 SEPLTGKLYQNLNLTINF--SNIFSPANTLEYISDNCPEKALNTGNCGLVE----- 160  
 QY 252 KHASALEEPFCGGTYNEEPERYSAISLPARKLRLGLTYIGKNELPIVFMFAAVLKLDL 311  
 DB 161 -----EKEDLIELDIS 205  
 QY 312 AMLDTHDCHMLTORCPNLEVLTRNV--IGDRGLEVLGRCKRLKRLIERGD-----DDQ 365  
 DB 366 GME--DEECTVS-----HRLIALS--QGCELEYMAVYVSDITNASLEHI--GTHLKNLCD 416  
 DB 206 GEGEFNADTLRLVSRNGLKELSDMGCTELSHITFTFN--LNCELDAMRALSLNNLPD 263  
 QY 417 FRLVLIDHEKITDPLDNGVRALLRGCDKRRFALYLRFG--GLTDVGLGYIGQYSPNVR 475  
 DB 264 LK-----DSDIELITCKFSKLAN--SLFLSKCIGLTDSLSLSTLTKLSQSIT 306

QY 476 WMLLGYGVE-SDAGLLEFAKGPSLQKLEMBGCLFFSERALAVAAATQLTSLRYL 528  
 DB 307 TLHGHGVEITDIGVQCILKSKRNTYIDFGCLRLSD-----IAVSAIAKLPYL 356  
 RESULT 8  
 FL2L RAT  
 ID FL2L RAT STANDARD; PRT; 276 AA.  
 AC Q9QZH7;  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE F-box/LRR-repeat protein 2-like.  
 GN FBL2.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A., INTERACTION WITH SKP1, SUBCELLULAR LOCATION, AND  
 RP TISSUE SPECIFICITY.  
 RC STRAIN=Buffalo; TISSUE=Liver;  
 RX MEDLINE=99439713; PubMed=10508920;  
 RA "Idin G.P., Rialland M., Glaiese D., Guguen-Guillouzo C.;  
 RT "Identification of a novel Skp2-like mammalian protein containing  
 RT F-box and leucine-rich repeats.";  
 RL FEBS Lett. 459:75-79(1999).  
 CC -!- FUNCTION: Probably recognizes and binds to some phosphorylated  
 CC proteins and promotes their ubiquitination and degradation.  
 CC -!- SUBUNIT: Interacts with SKP1. Part of a SCF (SKP1-cullin-F-box)  
 CC protein ligase complex (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- TISSUE SPECIFICITY: Widely expressed, with highest expression in  
 CC skeletal muscle, heart and brain.  
 CC -!- SIMILARITY: Contains 1 F-box domain.  
 CC -!- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; AF182443; AAF01221.1; -  
 DR PIR; T52349; T52349.  
 DR InterPro; IPR001810; F-box.  
 DR InterPro; IPR007089; LRR\_Cys.  
 DR InterPro; IPR006553; LRR\_cys\_sub.  
 DR Pfam; PF00646; F-box; 1.  
 DR SMART; SM00256; FBOX; 1.  
 DR SMART; SM00367; LRR\_CC; 1.  
 DR PROSITE; PS50181; FBOX; 1.  
 KW Ub1 conjugation pathway; Leucine-rich repeat; Repeat.  
 FT DOMAIN 22 68 F-BOX.  
 FT REPEAT 90 115 LRR 1.  
 FT REPEAT 142 167 LRR 2.  
 FT REPEAT 168 193 LRR 3.  
 FT REPEAT 194 219 LRR 4.  
 FT REPEAT 220 245 LRR 5.  
 FT REPEAT 246 271 LRR 6.  
 SQ SEQUENCE 276 AA; 30460 MW; 982994F91A265E1E CRC64;  
 Query Match 4.6%; Score 147.5; DB 1; Length 276;  
 Best Local Similarity 23.0%; Pred. NO. 0.00037;  
 Matches 70; Conservative 42; Mismatches 98; Indels 95; Gaps 13;  
 QY 233 NLVSVKLTDCBILDVNVFFKASA-----LSEFCGGTYNEEPERYSAISLPKLCGLGT 287  
 DB 56 NVLADGSGNWQRIDLFDQDIEGRVVENISKRCGGFLR-----KLSRGCL 102

QY 288 YIGKQELPIVPMFAVVKLLDLYAMLDTEHCHMLIQPCNLEVLTRNV--IGDRGLEV 345  
 Db 103 GVGDNALR---TFA-----QNCRNIEVLSUNGCTKTDTACTS 137  
 QY 346 LGRCKKRLKRLIRERGDDQGMEDGEGTVSHRGLIALSQGCELEYMAVTVSDITNASLE 405  
 Db 138 LSKFCSKLRHLDLASC-----SITNMSLKALSEGCPLEQLNI----- 176  
 QY 406 HIGTHLKNCLDFRLVLLDHEKIDTLPDNGVRALLRGCDKLRFAFYLRGGLT---DV 462  
 Db 177 -----SWCD-----QVT-----KGIQALVRGCGGLK--ALFLK--GCTOLEDE 211  
 QY 463 GLGVIGQYSPN-VRMMLLYGVESDAGLEFAKCPISLQKLEMGCLFFFSERALAVAAQ 521  
 Db 212 ALKYIGAHCPBLVTLNLQTCITQIDTEGLTTCRCGHKLQSLCAGSCNITDAILNALQON 271  
 QY 522 LLSLR 526  
 Db 272 CPLRLR 276

RESULT 9  
 ID\_FXL6 HUMAN STANDARD; PRT; 539 AA.  
 AC Q8NS31; Q8NS49; Q9UKC7;  
 DT 10-OCT-2003 (Rel. 42, last sequence update)  
 DT 15-MAR-2004 (Rel. 43, last annotation update)  
 DE F-box/LRR-repeat protein 6 (F-box and leucine-rich repeat protein 6).  
 DE FBXL6 OR PBL6.  
 GN FBXL6 (Human)  
 OS Homo sapiens (Human)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Skin;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.D., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullen S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield V.S.N., Krzywinski M.I., Skalek U., Smailus D.E., Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T., Suzuki Y., Oiyashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isegaki T., Sugano S.;  
 RT "NEDO human cDNA sequencing project.";  
 RN Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 90-372 FROM N.A. (ISOFORM 1/2).  
 RX MEDLINE=2003060; PubMed=10531035;  
 RA Cenciarelli C., Chiau D.S., Guardavaccaro D., Parks W., Vidal M., Pagano M.;  
 RT "Identification of a family of human F-box proteins.";  
 RL Curr. Biol. 9:1177-1179 (1999).

CC -!- FUNCTION: Probably recognizes and binds to some phosphorylated proteins and promotes their ubiquitination and degradation.  
 CC -!- SUBUNIT: Part of a SCF (SKP1-cullin-F-box) protein ligase complex (by similarity).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=Q8NS31-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q8NS31-2; Sequence=VSP\_008498;  
 CC Note=No experimental evidence available;  
 CC -!- SIMILARITY: Contains 1 F-box domain.  
 CC -!- SIMILARITY: Contains 3 leucine-rich (LRR) repeats.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; BC033066; AAH33066.1; -  
 CC EMBL; AK026541; BAB15499.1; -  
 CC EMBL; AF174592; AAF04513.1; -  
 CC Genbank; HGNC:13603; FBXL6.  
 CC InterPro; IPR001810; F-box.  
 CC InterPro; IPR001611; LRR.  
 CC InterPro; IPR006553; LRR cys sub.  
 CC InterPro; IPR008945; Skp1\_Skp2.  
 CC Pfam; PF00646; F-box; 1.  
 CC Pfam; PF00560; LRR; 1.  
 CC SMART; SM00367; LRR CC; 3.  
 CC PROSITE; PS50181; PROX; FALSE NEG.  
 CC Ubl conjugation pathway; Leucine-rich repeat; Repeat;  
 KW Alternative splicing.  
 FT DOMAIN 10 162 F-BOX.  
 FT REPEAT 219 244 LRR 1.  
 FT REPEAT 383 408 LRR 2.  
 FT REPEAT 497 528 LRR 3.  
 FT VARSPPLIC 208 213 Missing (in isoform 2).  
 FT /FTid=vsp\_008498.  
 SQ SEQUENCE 539 AA; 58587 MW; 8153185444653BB2 CRC64;  
 Query Match 4.4%; Score 141; DB 1; Length 539;  
 Best Local Similarity 21.1%; Pred. NO. 0.0029;  
 Matches 118; Conservative 62; Mismatches 162; Indels 218; Gaps 23;  
 QY 31 VVLDGVIPYIDDPKDRDAVQVCRWYELDSLTRKHVTIALCYTTTPARLRFRPHLESL 90  
 Db 128 VAADGPMFPL-----GRAARVCRWQEAASQP-----ALWHTVTLSPVGRPAKGV 175  
 QY 91 KLKGGPRAMFNLIPE-----DWGGHVTVPWKEISQYFDCLKSLHFRMIVKDS 139  
 Db 176 KAEKKLLASLEWMPNRSQQLRLTLHWKSVQHPVLVNG--ECCPLTF----- 224  
 QY 140 DLQNLARDRGVHLHALDKKSGFTTGLFHIGRCCKSLRVFLRES-----SILEK 191  
 Db 225 -----LKLGGCHGVTDALVLMKACCCQLHSLDLQHSVMVESVAVVSFLEE 269  
 QY 192 DGEWLEHLANNVTLETNLFYLTDAVVKIELELLAKNCPNLVSKLTDCEILDVLPNF 251  
 Db 270 AGSRMRKWLTVSSQTT-----AALG-----ALLGCCPQL-----QVLEV----- 305  
 QY 252 KHASALEEFCGGTYNEEPERYSAISLPKACLRGLTYIGKNELPIVFMFAVVKKLLDIY 311  
 Db 306 -----STGINRNSIPLQLPVEALQK----- 325  
 QY 312 AMLDTEHCHMLIQPCNLEVLTRNVIGDRGLEVLRGCCCKLRKLEIERGDDDDQGMEDDE 371  
 Db 326 -----GCPQLQVLRLLNL----- 345  
 QY 372 GTVSHRGLIALSQGCELEYMAVTVSDITNASLEHIGTHLKNCLDFRLVLLDHEKITDL 431



```

Db 346 G-----RG-VAPGPGFSPLEBELCLASSTCNFVSNEVLGRLLHGSPLRL----- 388
QY 432 PLONGVALLRGCDKLRFAFYLRGGGLTDVGLGYIGQYSPNVWMLLYGVGSDAGLLE 491
Db 389 -LD-----LRGCR-----ITPAGLDLPQRELEQ-----LHLGYGTSDBRLTLA 427
QY 492 -----FAKGCPSLOKLEMRGCLFRRERAL--AVAAQTQT-----SLRYLWVQGVGS 536
Db 428 KEGSPFTQKWCHTLRELDLSG-QGFSEKDLQALAAFLSTPGGSHPALCSNLNRTV 486
QY 537 PSGRDLVWARP-----FWNIE 553
Db 487 PSTVSSVISSCPGLLYLNLE 506

RESULT 10
ESAS TRYEQ STANDARD; PRT; 630 AA.
AC P2637;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative adenylate cyclase regulatory protein.
GN ESAG8C.
OS Trypanosoma equiperdum.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5694;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CL;
RX MEDLINE=91293076; PubMed=2065652;
RA Ross D.T., Raibaud A., Florent I.C., Sather S., Gross M.K.,
RT "The trypanosome VSG expression site encodes adenylate cyclase and a
RL leucine-rich putative regulatory gene.";
EMBO J. 10:2047-2053(1991).
CC -!- FUNCTION: May interact with adenylate cyclase to regulate its
CC activity.
CC -!- FUNCTION: May be involved in the posttranscriptional regulation of
CC genes in VSG expression sites.
CC -!- DEVELOPMENTAL STAGE: Expressed only in the bloodstream form of the
CC parasite.
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC -!- SIMILARITY: Contains 9 leucine-rich (LRR) repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X59385; CRA42028.1; -.
CC PIR; S16358; BWU78Q.
CC InterPro; IPR001611; LRR.
CC InterPro; IPR001841; Znf_ring.
CC Pfam; PF00560; LRR; 9.
CC Pfam; PF00097; ZF-C3HC4; 1.
CC SMART; SM00184; RING; 1.
CC PROSITE; PS00518; ZF_RING_1; 1.
CC PROSITE; PS00089; ZF_RING_2; 1.
KW Repeat; Leucine-rich repeat; cAMP biosynthesis; Zinc-finger;
KW DNA-binding.
FT ZN FING 10 46 RING-TYPE.
FT DOMAIN 47 54 ARG/LYS-RICH (BASIC).
FT DOMAIN 71 77 ARG/LYS-RICH (BASIC).
FT REPEAT 108 153 LRR 1.
FT REPEAT 181 225 LRR 2.
FT REPEAT 252 296 LRR 3.
FT REPEAT 322 366 LRR 4.
FT REPEAT 368 412 LRR 5.

```

```

FT REPEAT 414 458 LRR 6.
FT REPEAT 460 504 LRR 7.
FT REPEAT 506 550 LRR 8.
FT REPEAT 552 595 LRR 9.
SQ SEQUENCE 630 AA; 69989 MW; F882E49A6A6E6651 CRC64;

Query Match 4.4%; Score 139; DB 1; Length 630;
Best Local Similarity 20.9%; Pred. No. 0.0052;
Matches 110; Conservative 67; Mismatches 157; Indels 192; Gaps 29;

QY 126 LKSLHFRMIVKD-----SDLQNLARDGRGHVHLKLDKCSG 162
Db 161 LRKLRMKRTMVNDMWCSSISGLLKPLVHLEVDGGRGVTDTGLCLRK--TLEALSLSCLIN 218
QY 163 FTTDGLFHIGRFC--KSLRVLFLEESSILEKGEWHLHELALNNVTLETFN----YLTDI 216
Db 219 IITKG-----FDKICALPQLTSLSLCQTNVTKDLRCIHP-----DGKLKVLAYSSCHEITDL 270
QY 217 AVV-----KIDLEL-----LAKNC--PNLVSVKLTDCETL-----DLVNF----- 250
Db 271 TAICGMRSLKLSLSCGWNVTKLEELCKFSNRLRELDISGLVLGSAVLKLNILNKVL 330
QY 251 -----FKHASALEEF-----CGGTYNEBPBRYSAISLPKLCRLGLTYIGKNEL 294
Db 331 VSNCKNPKDLNGLERLVNLDKLNLSGCHGV-----SSLGFWANLSNLKELDISCES 382
QY 295 PIVFMFAAVLKLDLLYAMLDTHCMLIQRCPNLEVTETRNIVGDRGLVGLGCCRLK 354
Db 383 LVCFDGLQDLNNLEVLY-LRDVKSF-----TNVGAIKNLSKMRLELDLSG--CBRIT 430
QY 355 RLRIERDDDDQGMEDBEGTVSHRGLIALS-QGCSL-----HYMAVYVSDITN-- 401
Db 431 SL-----SGLETLKLEELSLEGGCEIMSPDPIWSLHLRLVLYVSECNLE 476
QY 402 -----ASLEHIGTH-----LKNLCDFRLVLDHHEKITDPLDNGVRAL 440
Db 477 DLGLEGITGLEELYHLGCKRCKTNFGPIWNLNRNVCVVELSCENLELDLSGLCLTGLEEL 536
QY 441 -LRGCDKLRRFALYLRGGGLTDVGLGYIGQYSPNV-----WMLLYGVGSDAGLLE 491
Db 537 YLIGCEI-----TPIGVVG-----NLRNLKCLSTCWC-----ANLKE 569
QY 492 FA--KGCPISQKLEMRGCLFRRERALAVAAQTQTSLRYL-WVQGVG 534
Db 570 LGGDLRLVNLKLDLSCCGGLSSVF---MELMSLPKQLQWFGYFG 611

RESULT 11
ESAS TRYBB STANDARD; PRT; 630 AA.
AC P23799;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative adenylate cyclase regulatory protein (Leucine repeat protein)
DE (VSG expression site-associated protein F14.9).
GN ESAG8 OR T-LR.
OS Trypanosoma brucei brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EATRO 164;
RX MEDLINE=91061751; PubMed=2247064;
RA Smiley B.L., Stadnyk A.W., Myler P.J., Stuart K.;
RT "The trypanosome leucine repeat gene in the variant surface
RT glycoprotein expression site encodes a putative metal-binding domain
RT and a region resembling protein-binding domains of yeast, Drosophila,
RT and mammalian proteins.";
RL Mol. Cell. Biol. 10:6436-6444(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=EATRO 1125;

```

RA MEDLINE=91081317; PubMed=2259625;  
RA Lips S., Revelard P., Pays E.;  
RT "A gene from the VSG expression site of Trypanosoma brucei encodes a  
RT protein with both leucine-rich repeats and a putative zinc finger.";  
RL Nucleic Acids Res. 18:7299-7303(1990).  
CC -!- FUNCTION: May interact with adenylate cyclase to regulate its  
CC activity.  
CC -!- FUNCTION: May be involved in the posttranscriptional regulation of  
CC genes in VSG expression sites.  
CC -!- DEVELOPMENTAL STAGE: Expressed only in the bloodstream form  
CC of the parasite.  
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.  
CC -!- SIMILARITY: Contains 23 leucine-rich (LRR) repeats.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
CC EMBL; M58701; AAA32117.2; ALT\_TERM.  
CC EMBL; X55978; CRA39448.1; -  
CC PIR; A36359; A36359.  
CC PIR; S13724; S13724.  
CC InterPro; IPR001611; LRR.  
CC InterPro; IPR001841; Znf\_ring.  
CC Pfam; PF00560; LRR; 8.  
CC Pfam; PF00097; zf-C3HC4; 1.  
CC SMART; SM00184; RING; 1.  
CC PROSITE; PS00518; ZF\_RING\_1; 1.  
CC PROSITE; PS00089; ZF\_RING\_2; 1.  
CC Repeat; Leucine-rich repeat; CAMP biosynthesis; Zinc-finger;  
KW DNA-binding.  
FT ZN FING 10 46 RING-TYPE.  
FT DOMAIN 47 54 ARG/LYS-RICH (BASIC).  
FT DOMAIN 71 77 ARG/LYS-RICH (BASIC).  
FT REPEAT 115 123 LRR 1.  
FT REPEAT 124 148 LRR 2.  
FT REPEAT 149 170 LRR 3.  
FT REPEAT 171 195 LRR 4.  
FT REPEAT 196 218 LRR 5.  
FT REPEAT 219 242 LRR 6.  
FT REPEAT 243 266 LRR 7.  
FT REPEAT 267 289 LRR 8.  
FT REPEAT 290 313 LRR 9.  
FT REPEAT 314 336 LRR 10.  
FT REPEAT 337 359 LRR 11.  
FT REPEAT 360 382 LRR 12.  
FT REPEAT 383 405 LRR 13.  
FT REPEAT 406 428 LRR 14.  
FT REPEAT 429 451 LRR 15.  
FT REPEAT 452 474 LRR 16.  
FT REPEAT 475 497 LRR 17.  
FT REPEAT 498 520 LRR 18.  
FT REPEAT 521 543 LRR 19.  
FT REPEAT 544 566 LRR 20.  
FT REPEAT 567 589 LRR 21.  
FT REPEAT 590 613 LRR 22.  
FT REPEAT 614 630 LRR 23.  
FT CONFLICT 22 22 V -> L (IN REF. 2).  
FT CONFLICT 25 25 L -> E (IN REF. 2).  
FT CONFLICT 38 38 Q -> E (IN REF. 2).  
FT CONFLICT 203 203 F -> C (IN REF. 2).  
FT CONFLICT 215 215 N -> S (IN REF. 2).  
FT CONFLICT 258 258 M -> P (IN REF. 2).  
FT CONFLICT 312 312 L -> V (IN REF. 2).  
FT CONFLICT 321 321 K -> R (IN REF. 2).  
FT CONFLICT 321 321 T -> S (IN REF. 2).  
FT CONFLICT 437 437 Y -> H (IN REF. 2).  
FT CONFLICT 462 462 Y -> H (IN REF. 2).  
FT CONFLICT 490 490 M -> L (IN REF. 2).  
FT CONFLICT 504 504 IWN -> FGI (IN REF. 2).

FT CONFLICT 511 511 C -> L (IN REF. 2).  
FT CONFLICT 522 522 D -> E (IN REF. 2).  
FT CONFLICT 543 546 EIT -> KIQP (IN REF. 2).  
FT CONFLICT 549 549 V -> I (IN REF. 2).  
FT CONFLICT 582 582 L -> V (IN REF. 2).  
FT CONFLICT 620 620 K -> E (IN REF. 2).  
SQ SEQUENCE 630 AA; 69998 MW; A65A35B5DCB50F7E CRC64;  
  
Query Match 4.2%; Score 133; DB 1; Length 630;  
Best Local Similarity 20.7%; Pred.No. 0.015;  
Matches 109; Conservative 69; Mismatches 156; Indels 192; Gaps 29;  
  
QY 126 LKSLHFRMIVKD-----SDQLNLPDRGHVHALKDKCSG 162  
DB 161 LRKLKMKRTWVNDWMCSSIGLLKFLVHLEVDGSRGVTDITGLFRLK--TLEALSLDNCIN 218  
QY 163 FTTDGLPHIGRFC--KSLRVLFLEESILKDGELHNLNNTVLETFN----YLTDI 216  
DB 219 ITK3-----FDKICALPOLTSLSLCQTNVTDKLRCHP-----DGKLMLODISSCHEITDL 270  
QY 217 AVV-----KIDLEL-----LAKNC--PNLVSVKLTDCIIL-----DLVNF----- 250  
DB 271 TAIGGVRSLEKLSLSCGWNVTKGLLEELCKPSNRLRELDISGCLVLSAVLKNLNLKVL 330  
QY 251 -----FKHASALEEF-----CGTYNEPERYSAISLPAKLCRLGLTIVIGNEL 294  
DB 331 VSNCKNFQDLNGLERLVNLEKLNLSGCHGV-----SSLGFVANLSNLKELDISGCES 382  
QY 295 PIVMFPAVLKGLDLYANLDTEDHCLMIQPCPNLEVTENVIGDRGLEVLGRCCRELK 354  
DB 383 LVCPDGLQDLNLEVLV-LRDVKSF-----TNVGAIKNLSKMRDLDSG--CERIT 430  
QY 355 RLRIERGDDQGMEDBEGTVSHRGLIALS--QGCSEL-----EYMAVYVSDITN-- 401  
DB 431 SL-----SGLETIKGLELSLEGCGEIMSPDPIWSLYHLRVLYVSECGNLE 476  
QY 402 -----ASLEHIGTH-----LKNICDFRLVLDHEEKITDPLNGVPAL 440  
DB 477 DLSGLQCLTGLEEMVYLGCRKCTNFGPIWNLNRLVLCVLELSCCNLDLDSGLQCLTGLEEL 536  
QY 441 -LRGCDKLRRFALYLRGGLTDVGLGYIGQVSPNVR-----WMLLYGVGESDAGLE 491  
DB 537 YLIGCEEI-----TTIGVVG-----NLNKLCLSTCWC-----ANLKE 569  
QY 492 FA--KGCPSLQKLEMRGCLFFSERALAVAATQLTSRLYL-WVOGYG 534  
DB 570 LGGHLRLVNLKLDLDSGCCGLSSSVF-----MELMSLPKLQWYFGF 611  
  
RESULT 12  
RAD7 YEAST  
ID RAD7 YEAST STANDARD; PRT; 565 AA.  
AC P06779;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE DNA repair protein RAD7.  
GN RAD7 OR YJR052W OR J1665.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP MEDLINE=90236305; PubMed=2158927;  
RA Melnick L., Sherman F.;  
RT "Nucleotide sequence of the COR region: a cluster of six genes in the  
RL yeast Saccharomyces cerevisiae.";  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87064430; PubMed=3023893;  
RA Perozzi G., Prakash S.;

RT "RAD7 gene of Saccharomyces cerevisiae: transcripts, nucleotide  
 sequence analysis, and functional relationship between the RAD7 and  
 RAD23 gene products.";  
 RL Mol. Cell. Biol. 6:1497-1507 (1986).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=S288C;  
 RA Huang M.-E., Manus V., Chuat J.-C., Galibert F.;  
 RT "Revised nucleotide sequence of the COR region of yeast Saccharomyces  
 cerevisiae chromosome X.";  
 RL Yeast 10:811-818 (1994).  
 CC -!- FUNCTION: This protein is one of 10 proteins (RAD1, 2,3,4,7,10,14,  
 16,23 and MMS19) involved in excision repair of DNA damaged with  
 UV light, bulky adducts, or cross-linking agents.  
 CC -!- MISCELLANEOUS: Mutants with mutations in the RAD7, RAD14, RAD16,  
 and RAD23 genes show partial incision defectiveness.  
 CC -!- SIMILARITY: TO S.POMBE SPC613.14.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M37696; AAB59347.1; -;  
 DR EMBL; M13015; AAB34953.1; -;  
 DR EMBL; L26347; AAB62860.1; -;  
 DR EMBL; L36344; AAB88755.1; -;  
 DR EMBL; Z49552; CAA89580.1; -;  
 DR PIR; A25226; A25226.  
 DR GeneOnline; 141885; -;  
 DR SGP; S0003813; RAD7.  
 DR GO; GO:0000113; C:nucleotide excision repair factor 4 complex; IDA.  
 DR GO; GO:000108; C:repairosome; IDA.  
 DR GO; GO:0003677; F:DNA binding; IDA.  
 DR GO; GO:0008094; F:DNA dependent ATPase activity; IDA.  
 DR InterPro; IPR007089; LRR\_cys.  
 DR DNA repair.  
 KW DOMAIN 1 200 HYDROPHILIC.  
 FT CONFLICT 278 279 LL -> FV (IN REF. 1).  
 FT CONFLICT 504 505 AC -> RP (IN REF. 1).  
 SQ SEQUENCE 565 AA; 63777 MW; 85C77357DC99737A CRC64;  
 Query Match 4.2%; Score 132; DB 1; Length 565;  
 Best Local Similarity 23.1%; Pred No. 0.016;  
 Matches 114; Conservative 78; Mismatches 188; Indels 114; Gaps 24;  
 QY 60 DSLTRKHVTIALCYTTTPARLRERPPHLESKLKGPRAAMNLPEDWGGHVTWVKEI 119  
 Db 131 ESLTKRQNTAKIQN---RRKKRAADLLDRRVNKVSSLSQSLCITKISENISKWQEA 187  
 QY 120 SQYFCLKSLHFRMI-----VKDSLQNLA---RDRGHVLHAKLKDCKSGFTTDLGFH 170  
 Db 188 ---DESSKLVFNKLRDLVGGVSTANLAKSKNALNDHTLQL----- 230  
 QY 171 IGRFCKS-LRVLFEESLSLEKDGWMLHNLNTVLTNPLYTDIAV-----VKIEDL 224  
 Db 231 ---FLKTDLRUTTFSDCKISFDG-----YKTLAIFSHLTSLQMGQQLNHSL 278  
 QY 225 ELLAKNCPNLVSKLTDCEILD---LVN-----FFKHASALEFCGGTYNEEPERSA 274  
 Db 279 LVIAEKLPLNKSIN-----LDGPFLLNEDTWKFPVINKGRLEFP---HISNTHRF 328  
 QY 275 ISLPAKLCRLGYTYIGKNELPIVFMAVLKLDLL--YAML-----DTEDHCLMILQPCP 327  
 Db 329 KSLSNLLNCGSTLVSLG-----LSRLDSISNYALLPQVLVDNDFHSLCLIEYFP 377  
 QY 328 NLEVLTRNVIGDGLGVLCRCCKLRKRIERGDDDDQGMEDDEEGTVSHRGLIALSQGCS 387  
 Db 378 NEEDVNDIEIIL-----NLIGQIGRTURKLVL-----NGCIDLTDSMINGLTAFIPEKC 426

QY 388 ELEYMAVYVSD-ITNASLEHI--GTHLKNL--CDFRLVLDHEEKITDLPDNGVRALLR 442  
 Db 427 PLEVLSLESDQITTSLSLTSFFSKVELNNLIESFRCLQGLDWAIIEILL-NGARDSLR 485  
 QY 443 GCDKLRFRFALYLRGGGLTDVGL-GYIGQYSPNVRMMLLGYVGSDDAGLLE-FAKGCPSLQ 500  
 Db 486 S-----LNLNSLKELTKFAFVALACPNLTLYLDLGFVRCVDDSVIQLMGEGQNPFLT 535  
 QY 501 KLEMRGCLFFSERA 514  
 Db 536 VIDVFGDNLVTEKA 549  
 RESULT 13  
 NA12 HUMAN STANDARD; PRT; 1061 AA.  
 ID NA12\_HUMAN AC P59046;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE NACHI-, LRR- and PYD-containing protein 12 (PYRIN-containing APAF1-  
 like protein 7) (Monarch-1).  
 GN NALP12 OR PYPAF7.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RN SEQUENCE FROM N.A. (ISOFORM 1).  
 RP MEDLINE=22162427; PubMed=12019269;  
 RA Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,  
 RA Lora J.M., Geddes B.J., Briskin M., DiStefano P.S., Bertin J.;  
 RT "PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates  
 activation of NF-kappa B and caspase-1-dependent cytokine  
 processing.";  
 RT J. Biol. Chem. 277:29874-29880 (2002).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=22451042; PubMed=12563287;  
 RA Teschopp J., Martinon F., Burns K.;  
 RT "NALPs: a novel protein family involved in inflammation.";  
 RL Nat. Rev. Mol. Cell Biol. 4:95-104 (2003).  
 RN [3]  
 RN SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).  
 RC TISSUE=Lymphoma;  
 RA Williams K.L., Linhoff M.W., Harton J.A., Ting J.P.Y.;  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RN SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Leukocyte;  
 RX MEDLINE=2238257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Millatny S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Truchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Goughwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 CC -!- FUNCTION: May mediate activation of CASP1 via ASC and promote



induces glucose repression. Is not an essential protein. Involved in substrate recognition in ubiquitin-dependent degradation.

-1- PATHWAY: Necessary for the glucose repression pathway.

-1- SUBUNIT: Part of a SCF E3 ubiquitin ligase complex containing CBP3, CDC53, HRT1 and GRI1.

-1- SUBCELLULAR LOCATION: Associated with the particulate fraction. Probably forms a complex by protein-protein interactions via its leucine-rich segment.

-1- INDUCTION: Expressed constitutively at low levels.

-1- SIMILARITY: Contains 1 F-box domain.

-1- SIMILARITY: Contains 12 leucine-rich (LRR) repeats.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; M59247; AAA34652.1; -  
EMBL; Z49590; CAA89617.1; -  
EMBL; I47993; AAB39313.1; -  
PIR; A41529; A41529.  
GeneOnline; 141922; -  
SGD; S0003850; GRI1.  
GO; GO:0000142; C:contractile ring (sensu Saccharomycetes); IDA.  
GO; GO:0005737; C:cytoplasm; IDA.  
GO; GO:0005634; C:nucleus; IDA.  
InterPro; IPR001810; F-box.  
InterPro; IPR001611; LRR.  
InterPro; IPR007089; LRR\_cys.  
Pfam; PF00646; F-box; 1.  
Pfam; PF00560; LRR; 1.  
SMART; SM00256; FBOX; 1.  
PROSITE; PS50181; FBOX; 1.  
Glucose metabolism; Ubl conjugation pathway; Repeat;  
Leucine-rich repeat.  
DOMAIN 314 361 F-BOX.  
REPEAT 409 434 LRR 1.  
REPEAT 435 460 LRR 2.  
REPEAT 461 486 LRR 3.  
REPEAT 487 512 LRR 4.  
REPEAT 513 548 LRR 5.  
REPEAT 549 567 LRR 6.  
REPEAT 568 593 LRR 7.  
REPEAT 594 619 LRR 8.  
REPEAT 620 645 LRR 9.  
REPEAT 646 670 LRR 10.  
REPEAT 671 699 LRR 11.  
REPEAT 700 726 LRR 12.  
POLY-ASN.  
DOMAIN 1045 1124 ASN-RICH.  
SEQUENCE 1151 AA; 132733 MW; 6BB6C46611B6F825 CRC64;

Query Match 4.0%; Score 127.5; DB 1; Length 1151;  
Best Local Similarity 20.7%; Pred. No. 0.091;  
Matches 119; Conservative 78; Mismatches 190; Indels 189; Gaps 29;

QY 11 LSLRSMTEENVKTRVVDVLDVIPPY-----IDDPKRDV----- 49  
DB 212 MRLRALETENMEIRNLRL--KILTIIEEYKSLYAYCHSKLRGQQVENPTDNFIWINSI 269  
QY 50 -----SQVCRWVELDSLTKRHVTIAL-----CYTTTPARLRRFP----- 85  
DB 270 DTTESSDLKEGLQDLSEYGRQFINNVLSNPSNQICTSVT-----RRSPFALNMLPSRI 324  
QY 86 -HLESKLKXKPRAAFNIPEDWGGHTVPWKVEISQYFDCUKSLHFRMIVKDSLQNL 144  
DB 325 LHILDKLNQKYDIVKFLTVSKLWA-----EI-----IVKILYRPHINKSKQLDLF 371  
QY 145 ARDRGHVHALKDKCSGGFTDGLFHIGRPFCKSLRVLFLSESSILEKDGWLHELALNNT 204

372 -----LRMKLT-----SEETVFNRLMKRLNFSFV-----GDYMHDTELNYF 410  
QY 205 VLETLNLYLTDIAVVKIEDLELLAKCNPNLVSKLTDCEILDVNFPHASALEEFCGGT 264  
DB 411 V-----GCKNL-----ERLTIV-FCKHTSV----- 430  
QY 265 YNEEPERYSAISLPKLCRLGLTVIGKNELPIVFEMFAAVLKLDLVLAMLTDEHCHM--L 322  
DB 431 -----PISAVLRGCKF-----LQSVDT-GIRVSDVDVDTL 461  
QY 323 IQRCPNLE-----VLETRNVIGD--RGLEVLGRCCRLKRLRIERGGDDQGDDESGTVSHR 377  
DB 462 ATVCPRVQGVFPQARNVTFDSLRFVHS--PMLKRIKITANN--MNDE----- 508  
QY 378 GLI-ALSQGSELEYMAVYVS-DITNASLEHGHGTHLKNLQDFRLVLDLHBEKITDPLDN 435  
DB 509 -LVELLANKPFLLEVDTITLSPNVTSSLLKLTRLVQLREFRIT--HNTNIT-----DN 560  
QY 436 GVRALLRGCDKLRFPALYLRG--GLTGVIGYIGQYSPNVRMMLLGYVG-SDAGLEF 492  
DB 561 LFQELSKVDDMPSLRLLDLSGCENITDKTIESIVNLAPKLRNVFLGKCSRITDASLFOL 620  
QY 493 AKGCPSLQKLEMRGCLFSPERALAATAATQTLSTLYL 528  
DB 621 SKLGNLQTVHFGHCFNITDNGVRLFHSCTRIQYV 656

RESULT 15  
FXL4 HUMAN  
ID FXL4 HUMAN STANDARD; PRT; 621 AA.  
AC Q9UKA2; O95919; Q9UJU0.  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE F-box/LRR-repeat protein 4 (F-box and leucine-rich repeat protein 4)  
DE (F-box protein FBL5).  
GN FBL4 OR FBL4 OR FBL5.  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND SUBCELLULAR LOCATION.  
RX MEDLINE=2003061; PubMed=10531037;  
RA Winston J.T., Koepf D.M., Zhu C., Elledge S.J., Harper J.W.;  
RT "A family of mammalian F-box proteins."  
RL Curr. Biol. 9:1180-1182 (1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=2003060; PubMed=10531035;  
RA Cenciarelli C., Chiar D.S., Guardavaccaro D., Parks W., Vidal M.,  
RA Pagano M.;  
RT "Identification of a family of human F-box proteins."  
RL Curr. Biol. 9:1177-1179 (1999).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20399565; PubMed=10945468;  
RA Ilyin G.P., Rialland M., Pigeon C., Gugen-Guillouzo C.;  
RT "cDNA cloning and expression analysis of new members of the mammalian F-box protein family."  
RL Genomics 67:40-47 (2000).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX Palmer S.;  
RL Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Eye;  
RX MEDLINE=22398257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
CC -!- SUBUNIT: Part of a SCF (SKP1-cullin-F-box) protein ligase complex  
CC (By similarity).  
CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.  
CC -!- TISSUE SPECIFICITY: Expressed in heart, kidney, liver, lung,  
CC pancreas, and placenta, but not in skeletal muscle.  
CC -!- SIMILARITY: Contains 1 F-box domain.  
CC -!- SIMILARITY: Contains 8 leucine-rich (LRR) repeats.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC  
CC EMBL; AF176699; AAF03699.1; -  
CC EMBL; AF174590; AAF04511.1; -  
CC EMBL; AF199395; AAF09247.1; -  
CC EMBL; AL022395; CAB37981.1; -  
CC EMBL; BC055010; AAH55010.1; -  
CC EMBL; HGNC:13601; FBXL4.  
CC MIM; 605654; -  
CC InterPro; IPR001810; F-box.  
CC InterPro; IPR007089; LRR\_cys.  
CC InterPro; IPR006553; LRR\_cys\_sub.  
CC InterPro; IPR008945; Skp1\_Skp2.  
CC Pfam; PF00646; F-box; 1.  
CC SMART; SM00256; FBOX; 1.  
CC SMART; SM00367; LRR\_CC; 1.  
CC PROSITE; PS0181; FBOX; 1.  
CC Ubl conjugation pathway; Repeat; Leucine-rich repeat; Nuclear protein.  
KW DOMAIN 277 332 F-BOX.  
FT REPEAT 400 421 LRR 1.  
FT REPEAT 425 449 LRR 2.  
FT REPEAT 450 475 LRR 3.  
FT REPEAT 478 503 LRR 4.  
FT REPEAT 504 524 LRR 5.  
FT REPEAT 532 558 LRR 6.  
FT REPEAT 559 583 LRR 7.  
FT REPEAT 584 609 LRR 8.  
FT CONFLICT 5 5 P -> Y (IN REF. 1).  
FT CONFLICT 42 42 S -> G (IN REF. 3).  
FT CONFLICT 122 122 Q -> K (IN REF. 1).  
FT CONFLICT 126 126 Y -> I (IN REF. 1).  
FT CONFLICT 219 219 E -> G (IN REF. 1).  
FT CONFLICT 232 232 P -> A (IN REF. 1).  
SQ SEQUENCE 621 AA; 70096 MW; D8FD51A5C4F922D3 CRC64;  
Query Match 4.0%; Score 126.5; DB 1; Length 621;  
Best Local Similarity 21.4%; Pred. No. 0.047;  
Matches 98; Conservative 65; Mismatches 161; Indels 135; Gaps 23;  
QY 121 QYFDCKLSLHFR----RMIVKDSLDLQNLARDRGHVLKLDKCSGGTTDGLFHIGRFCK 176  
DB 190 QFKPKQIKQINFPFTNIRLEVNSSLLEYITELDAVVLHGKV-DKPVLSLKTSLIDMND--- 245

QY 177 SLRVLFERSIILEKDGEMHEL-----ALNNTVLETNLFYLTDAV--VKIED 223  
DB 246 -----IEDDAYAEKDCGMDSLNKKFSSAVLGGPNNGYFDKLPYELIQLIINHLLTPD 299  
QY 224 LELAKNCPNLVSVKLTDCETLDLVNFFKHASALEEFCCGGTYNEEPERYSAISLP---AK 280  
DB 300 LCRUAQTC-----KLLSQHCC-----DPLQYIHLNLQPYWAK 331  
QY 281 L-----CR-----LGYTYGKNELPIVFNFAVLK-----KLDLYLAMLDTED 318  
DB 332 LDDTSLEFLQSRCTLVQWMLNLSWTNCRGFIISVAGFSRFLKVCSELVRLKELUSCHFLNET 391  
QY 319 HCM--LIQRCNLEVLTRNV--IGDRGLEVLGRCCCKELKELRIERGDDDDQDMEDEEQT 374  
DB 392 -CLEVISEMCPNLQALNLSSCDKLPQAFNHIKLS--LKLVLVYR-----TKV 438  
QY 375 SHRGILIALSQGCELEYMA---VYVSDITNASLEHIGTHLKNLCDFRLVLLDHEEKTD 430  
DB 439 EQTALLSILNFCSELQHLSLGSCVMIEDY-DVIASMIKAKCKKL---RTLDLWRCKNIT- 493  
QY 431 LPLDNGVRLLRGGCDKLRFPALYLRGGLTVDGLGY-----IGQYSPNVNRM 477  
DB 494 ---ENGIAELASGCPLE-----ELDLGMCPTLOSSTGCTFRLAHQLNLOKL 538  
QY 478 LL-GYVGESDAGLLEFAKGCPSLOKLEMRGCLFFSERAL 515  
DB 539 FLTANRSVCDIDIDELACNCTRLOQLDILGTMVSPASL 577  
Search completed: April 20, 2004, 01:36:59  
Job time : 22 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2004, 01:09:15 ; Search time 28 seconds  
(without alignments)  
2081.859 Million cell updates/sec

Title: US-10-009-791-22

Perfect score: 3180

Sequence: 1 TKTSAPFLTLRLSNNTTE.....QRSDFPDVVPDLATATCVDT 606

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_78:\*  
1: Pirl1.\*  
2: Pirl2.\*  
3: Pirl3.\*  
4: Pirl4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2126.5	66.9	592	T52139	LRR-containing F-b
2	818	25.7	585	E85040	hypothetical prote
3	812.5	25.6	594	T48087	transport inhibito
4	749.5	23.6	577	F86261	F13K23.7 protein -
5	693	21.8	614	T09902	hypothetical prote
6	229.5	7.2	618	T48193	hypothetical prote
7	209.5	6.6	712	C71419	hypothetical prote
8	188	5.9	607	E96598	protein F20N2.2 [i
9	187.5	5.9	628	E84649	probable glucose r
10	186	5.8	489	S44609	hypothetical prote
11	176.5	5.6	467	G96837	unknown protein T2
12	176.5	5.6	690	T08604	hypothetical prote
13	172.5	5.4	518	D96512	hypothetical prote
14	169.5	5.3	563	T41312	probable DNA excis
15	168.5	5.3	522	T45861	hypothetical prote
16	159.5	5.0	656	T84547	hypothetical prote
17	154	4.8	463	T39987	probable regulator
18	152	4.8	381	T43444	hypothetical prote
19	149.5	4.7	678	B84856	hypothetical prote
20	148	4.7	554	B85072	hypothetical prote
21	147.5	4.6	276	T52349	F-box protein FB12
22	147.5	4.6	368	F86291	hypothetical prote
23	146	4.6	360	A96799	hypothetical prote
24	140	4.4	360	B6347	hypothetical prote
25	139	4.4	630	BWTT8Q	regulatory protein
26	139	4.4	630	S13724	ESAG 8 protein - T
27	135.5	4.3	934	T05201	hypothetical prote
28	134.5	4.2	1184	T71436	hypothetical prote
29	134.5	4.2	1301	D85188	disease resistance

30 133 4.2 630 2 A36359  
31 132 4.2 565 2 A25226  
32 131.5 4.1 301 2 E85358  
33 131.5 4.1 1217 2 T52348  
34 127.5 4.0 1151 2 A41529  
35 126.5 4.0 1405 2 T04426  
36 126 4.0 551 2 T23345  
37 125 3.9 307 2 A85069  
38 124.5 3.9 250 2 T08680  
39 122.5 3.9 1778 2 AF1116  
40 122 3.8 1215 2 H84513  
41 121.5 3.8 556 2 T33367  
42 121.5 3.8 578 2 E96838  
43 121 3.8 465 2 D96567  
44 120.5 3.8 1189 2 T52346  
45 119.5 3.8 419 2 H96695

#### ALIGNMENTS

##### RESULT 1

T52139  
LRR-containing F-box protein [imported] - Arabidopsis thaliana  
N:Alternate names: hypothetical protein At2g39940  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 16-Feb-2001  
C:Accession: T52139; C84823  
R:Xie, D.X.; Reys, B.F.; James, S.; Nieto-Rostro, M.; Turner, J.G.  
Science 280, 1091-1094, 1998  
A:Title: Coll: An Arabidopsis gene required for jasmonate-regulated defense and fertility  
A:Reference number: Z25980; MUID:98248619; PMID:9582125  
A:Accession: T52139  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-592 <X1E>  
A:Cross-references: EMBL:AF036340; PIDN:AAC17498.1  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L. euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: C84823  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-592 <STO>  
A:Cross-references: GB:AB002093; NID:g2088647; PIDN:AAB95279.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: Coll: At2g39940  
A:Map position: 2

Query Match 66.9%; Score 2126.5; DB 2; Length 592;

Best Local Similarity 69.6%; Pred. NO. 6.9e-159;

Matches 407; Conservative 72; Mismatches 97; Indels 9; Gaps 3;

QY 19 EERNVKTIV-----VQVWLCVTPYDDPKDRAVSQVCRWYELDSITKQHTIALCY 73

Db 2 EDDPIKCKLSVATVDDVIRQVMTYITDPKDRDSASLVCRWFKIDSETREHVTMALCY 61

QY 74 TTTTARLRRRPHLESLLKKGKPRAAFNLPEDWGGHVTWVKESIQYFQCLKSLHFR 133

Db 62 TATPDRSRRRPNRLSKLKGKPRAAFNLPENWGGVTPWVTETISNNLQKLSVHFR 121

QY 134 MIVKDSLDLQNLARDGRGHVTHALKDKCSGFTTGDGLFHIGRCKSLRVLFLESSILEKDG 193

Db 122 MIVSDLLDLRAKARADDETLKDKCSGFTTGDGLLSIVTHCRKIKTLMESSSFEXDG 181

QY 194 EWLHELALNNVTFLNFYLTDAIVVKTEDELELANCPNLVSVKLTDCETLIDLNVFKH 253

Db 182 KWLHQLAQNHTSLEVLNFYMTFAKISPKOLETTIARNCRSLVSVKVGDFEILELVGFKA 241

QY 254 ASALBEFCGGTYNEE---PERYSALSPLAKLRLGLTYIGKNELPIVFMFAVLKLDLL 310

```

Db 242 AANLEPCGSLNEDIQMPKYMVLVPRKLCRLGLSYMGNEPILFFPAAQIRKLDLL 301
QY 311 YAMLDTHDHCWLIQCPNLEVLTRNIVIGRGLVGRCCCKRLKRLRIERGDDDDQMEDE 370
Db 302 YALLEDETHCTLIQCPNLEVLTRNIVIGRGLVGRCCCKRLKRLRIERGADQMEDE 361
QY 371 EGVSHRGLIALSGCSELEFMAVYVSDITNASLEHIGTHLKNICDFRLVLLDHEEKITD 430
Db 362 EGLVSQGLIALAQCGELETMAVYVSDITNESLESIGTYLKNICDFRLVLLDHEEKITD 421
QY 431 LPLDNGVRAILLGCDKLRRLPALYLRGGITDVLGYIGQYSPNVRWMLLGVVGSDEGLL 490
Db 422 LPLDNGVRSLLIGCKLRRLPALYLRGGITDVLGYIGQYSPNVRWMLLGVVGSDEGLM 481
QY 491 EFAGKCPSLQKLEMGCLFFSFERALAAVAATQTSRLYLWVGYSVPSGRDLLVWAPFFW 550
Db 482 EFSRGCPNLQKLEMGCC-FSERALAAVAATQTSRLYLWVGYSVPSGRDLLVWAPFFW 540
QY 551 NIELPSKXVATNNPDDTVVVEHPAHILAYISLAGQSDPDDTV 595
Db 541 NIELPSRRVENVQCGEIREMEHPAHILAYISLAGQSDPDDTV 585

RESULT 2
E85040
hypothetical protein AT4903190 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C:Accession: E85040
R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617199
A:Accession: E85040
A:Status: Preliminary
A:Molecule type: DNA
A:Cross-references: GB:NC_001268; NID:97270189; PIDN:CAB77804.1; GSPDB:GN00140
A:Residues: 1-585 <STO>
A:Gene: AT4903190
A:Map position: 4

Query Match 25.7%; Score 818; DB 2; Length 585;
Best Local Similarity 34.6%; Pred. No. 3.8e-56;
Matches 200; Conservative 109; Mismatches 237; Indels 32; Gaps 12;

QY 32 VLDCVPIYDDPKORDAVSQVCRWYELDSLTKRHVTIALCYTTTPARLRPPHLESUK 91
Db 9 VLEHILSFIDSNEGRNSVSVCKSWFETERKTRKRVFVGCYAVSPAATRRFPPEMRSLT 68
QY 92 LKGKPRAMFNLIPEDWGHHVTPWKEISQYFDCLSKLHFRMIVKDSLOQLA---RDR 148
Db 69 LKGKPHFADNLPDGGWGYAMPWEIEMAAKSSSEIRMKRMVTVDECLEKTAASFQD- 127
QY 149 GHVLHAKLDCSGFTDGLPHIGRCKSLRVLFLEESSILEKDGWELHNLANNVTLET 208
Db 128 ---FKVLVLTSCRGFTDGLAAIAATCRNLKELDLRESVDVDSVGHLSHPDDTVTSVSLNI 184
QY 209 LNFVLTDAVVKIELELAKNCPNLVSVKLTDCETIID-LVNFHFKASALEEFCGGTYNE 267
Db 185 LDFSCLD-SEVKISDLERLSRSPNLKSLNPAVTLGLVSLRCAPQLTELGTGSFAA 243
QY 268 E--PERSAISLPAKLCR-----LGLTYIGKNELPIVEMFAAVLKLDLILYAMLDTEHCH 320
Db 244 QLKPEAPSKLSEAFNSCKQLQSLGLMDVLPEYLPALYVCPGLTSLNLSYATVRPMDLV 303
QY 321 MLQRCPNLEVLTRNIVIGRGLVGRCCCKRLKRLRIERGDDDDQMEDEEGTVSHRGLI 380
Db 304 ELURRCKIQLKMWMDLIEDKGLNAVASYCKELRELVFSEPD--LDATNIPTEGLV 361
QY 381 ALSQGCSELEFMAVYVSDITNASLEHIGTHLKNICDFRLVLLD--HEEKITDPLDNGVR 438

```

```

Db 362 FVSKGCRKLESVLYFCVQFTNAALFTIARKRPNLKCFRLCVIEFPADPYKTNPDLKGF 421
QY 439 ALARGCDKLRRLPALYLRGGITDVLGYIGQYSPNVRWMLLGVVGSDEGLLFAKGCPS 498
Db 422 AIAEGGRDLRLSV---SGLSDKAFYIGKHAKYRMLSLAFAGDSIDLMLHLLSGCES 478
QY 499 LQKLEMGCLFFSFERALAAVAATQTSRLYLWVGYSVPSGRDLLVWAPFFWNIELPSR 558
Db 479 LKLEIRDCP-FGDTALLEHAALKLETMRSLWMSSCFVSGACKLLSQMPLNAVTEVI--- 534
QY 559 KVATNNPDDTVVVEHPA-HILAYISLAGQSDPDDTV 595
Db 535 ----DEHPSPRESSPVERIYIYRTVAGPRMDTPEFV 568

RESULT 3
T48087
transport inhibitor response protein TIR1 [imported] - Arabidopsis thaliana
N:Alternate names: protein T20010.80
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 03-Nov-2000
C:Accession: T48087; T51946
R:Obermaier, B.; Ottenwaelder, B.; Duchemin, D.; Zeitler, K.; Mewes, H.W.; Rudd, S.; Le
submitted to the Protein Sequence Database, April 2000
A:Reference number: Z24484
A:Accession: T48087
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-594 <OBE>
A:Cross-references: EMBL:AJ163816
A:Experimental source: cultivar Columbia; BAC clone T20010
R:Ruegger, M.; Dewey, E.; Gray, B.; Hobbie, L.; Turner, J.; Estelle, M.
submitted to the EMBL Data Library, May 1997
A:Description: The TIR1 protein of Arabidopsis functions in auxin response and is relat
A:Reference number: Z25881
A:Accession: T51946
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-594 <RUE>
A:Cross-references: EMBL:AF005047; PIDN:AA69175.1
C:Genetics:
A:Gene: TIR1
A:Map position: 3
A:Introns: 156/2; 320/3
A:Note: T20010.80

Query Match 25.6%; Score 812.5; DB 2; Length 594;
Best Local Similarity 34.8%; Pred. No. 1e-55;
Matches 201; Conservative 104; Mismatches 250; Indels 23; Gaps 11;

QY 32 VLDCVPIYDDPKORDAVSQVCRWYELDSLTKRHVTIALCYTTTPARLRPPHLESUK 91
Db 13 VLEHVSFIQDKDRNSVSVCKSWFETERKTRKRVFVGCYAVSPAATVIRFPKVSVE 72
QY 92 LKGKPRAMFNLIPEDWGHHVTPWKEISQYFDCLSKLHFRMIVKDSLOQLAARDGHV 151
Db 73 LKGKPHFADNLPDGGWGYVYPTTEAMSSSYTVEELRLKRMVTVDDCLELIAKSPKN- 131
QY 152 LHALKLDKCSGFTDGLPHIGRCKSLRVLFLEESSILEKDGWELHNLANNVTLETNLF 211
Db 132 FKVLVLTSCRGFTDGLAAIAATCRNLKELDLRESVDVDSVGHLSHPDDTVTSVSLNI 191
QY 212 YLTDAVVKIELELAKNCPNLVSVKLTDCETIID-LVNFHFKASALEEFCGGTYNE-- 268
Db 192 SCL-ASEVSFSALERLVTRCPNLKSLNRAVPEKLAATLLQAPOLESLGTGGYTAERV 250
QY 269 PERSAISLPAKLCR-----LGLTYIGKNELPIVEMFAAVLKLDLILYAMLDTEHCH 323
Db 251 PDVYGLSVALSGCKELRCLSGFDWAVPAYLVAVSVCSRLTTLNLSYATVQSYDLVKLL 310
QY 324 QRCPNLEVLTRNIVIGRGLVGRCCCKRLKRLRIERGDDDDQMEDEEGTVSHRGLIALS 383
Db 311 CQCPKLRQLWLDYIEDAGLEVASTCKDLRLRV--FFSEPFVMEPNVALTEQGLVS 368

```







A:Accession: E96598  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-607 <STO>  
 A:Cross-references: GB:AE005173; NID:G8778504; PIDN:AAF79512.1; GSPDB:GN00141  
 C:Genetics:  
 A:Gene: P20N2.2  
 A:Map position: 1

Query Match 5.9%; Score 188; DB 2; Length 607;  
 Best Local Similarity 23.0%; Pred. No. 1.2e-06;  
 Matches 132; Conservative 74; Mismatches 185; Indels 184; Gaps 27;

QY 59 LDSLIRKHVTTALCVTT--PARLRPPHLESKLKGPRAAMFNLPEDWG-GHVTWP 115  
 DB 14 LKRLDLENLCVACVSTTIRGAVGVLPSTSLD-----SVFS--PDDETILNHVLRG 65  
 QY 116 VKEISQY--FDGLKSLHFRMTVKDSLQNLARDGHV---LHAKLKDCSGFTTDLGLPHI 171  
 DB 66 CIGUSSLTNLCR-----LNAASVRGVJGPHURELHLRCSLLSTVLTLYI 111  
 QY 172 GRFCKSRVLEBESSILEK-----GEWLHELALN-----NTVLETLNFI 212  
 DB 112 GTLCPLNRLVLEWADLSDPDVFSQNLTMQLNGCPYLESIQINIRGILVDATAFQSVRF 171  
 QY 213 LTDIAVVKIEDELELLAKCPNLVSKLTDCEILDVNFPKHASALEBFCGGTYNEEPERY 272  
 DB 172 LP-----ETIKALRLQ-----PLLESAILLMNRFK-----VTGTYLSQPDYN 209  
 QY 273 SAISLPA-----KLCRLGLTVIGKNELPIVFMAFAVLKLL-----DLLYAM 313  
 DB 210 SALLSPSPFTLQSUSVLDLISRLIIAIGTSPQLVKLDLEDRPEKPFDDNDLTYTG 269  
 QY 314 LDTEDHCH-----LIQCPNLEVLNTRNVIGDRGLEVLGRCCKRLKRLRIERGDDQGM 367  
 DB 270 LQALGFCCQLTSLSVTRCYNRKISFKR--INDMGIPLLSACKGLESVRL----- 318  
 QY 368 EDEEG---TVSHRGLIALSQSCSELYM-----AVTVSD----- 398  
 DB 319 ----GGFPKVDAGPASPALLSHCRNLKPFVRGAFLLSDLAFLHDVTGSSCSLQEVRLSTCP 374  
 QY 399 -ITNASLEHIGT-----HLKNCDFRLVLLDHEEKITDPL-----DNGVRAL- 440  
 DB 375 LITSEAVKGLGCGNLEVLGSCCKSIDSCNSVSAURKLTSLNLAGADVTDSGLALG 434  
 QY 441 -----LRGCDKLRFRFALYLRGGITDVLGYI----GOYSPNVNRMMLGYV-GES 485  
 DB 435 KSDVPIITQLSLRG-----RR--VSDRGISYLLNNEGTTISKTSLTDLGHMPCIS 482  
 QY 486 DAGLLEFAKGCPSLOKLEMRGCLFFSERALAVAT 520  
 DB 483 DRAIHTITHCCKALTELSIRSCFHVTDSIESLAT 517

RESULT 9  
 A84649  
 probable glucose regulated repressor protein [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C:Accession: A84649  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vankken, S.E.; Umayam, L.; Tallon, L.;  
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; MUID:20083487; PMID:10617197  
 A:Accession: A84649  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-628 <STO>  
 A:Cross-references: GB:AE002093; NID:G4432860; PIDN:AAD20708.1; GSPDB:GN00139  
 C:Genetics:  
 A:Gene: At2G25490

A:Map position: 2

Query Match 5.9%; Score 187.5; DB 2; Length 628;  
 Best Local Similarity 25.6%; Pred. No. 1.3e-06;  
 Matches 77; Conservative 43; Mismatches 102; Indels 79; Gaps 11;

QY 220 KIEDLEL--LAKNCPNLVSVKL-----TDCIILDVNVFFKHASALEBFCGGTYNEEPER 271  
 DB 163 KVSGLGRSIRGSCPSLGSLSLNMVSTITDNGLEIA-----EGCAOLEKLELNLR 212  
 QY 272 YSAISLPAKLCRLGLTVIGKNELPIVFMAFAVLKLLDLYAMLTDHCLMIQRCPNLE- 330  
 DB 213 CSTIT-----DKGLVAIAKS-----CPNLTE 233  
 QY 331 -VLTRNVIGDRGLEVLGRCCKRLKRLRIERGD--DQGMDEBEGTVSHRGLIALSQCS 387  
 DB 234 LTLEACSRIGDEGLLAIRASCCLKSVSKNCPLVRDQGIAS-----LLSNTTC 282  
 QY 388 ELEYMAVYVSDITNASLEHIGTHLKNLCDPFLVLLDHEEKITDPLDNGVRALLRGCDKL 447  
 DB 283 SLAKLKLQMLNVTVDVSLAVVGHYGLSITDLVLGLSHVSEKGFVWGMGV-----GLOKL 337  
 QY 448 RRFALYLRGGITDVLGYIGQYSPNVNRMMLGYVG-ESDAGLLEFAKGCPSLOKLEMRG 506  
 DB 338 NSLTITACQ-GVTDMGLESVKGCPNKKAIISKSPILLSONGLVSAKASLSLSLQLEE 396  
 QY 507 C 507  
 DB 397 C 397

RESULT 10  
 S44609  
 hypothetical protein C02F5.7 - Caenorhabditis elegans  
 N:Contains: hypothetical protein C30A5.1  
 C:Species: Caenorhabditis elegans  
 C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Sep-1997  
 C:Accession: S44609; S44774  
 R:Anderson, K.  
 submitted to the EMBL Data Library, May 1993  
 A:Description: Sequence of the C. elegans cosmid C02F5.  
 A:Reference number: S44603  
 A:Accession: S44609  
 A:Molecule type: DNA  
 A:Residues: 1-489 <AND>  
 A:Cross-references: EMBL:L14745; NID:G289607; PID:G289614  
 A:Experimental source: cosmid C02F5  
 R:Anderson, K.  
 submitted to the EMBL Data Library, February 1993  
 A:Description: Sequence of the C. elegans cosmid C30A5.  
 A:Reference number: S44774  
 A:Accession: S44774  
 A:Molecule type: DNA  
 A:Residues: 1-197 <ANW>  
 A:Cross-references: EMBL:L10990; NID:G156211; PID:G156212  
 A:Experimental source: cosmid C30A5  
 A:Note: designated as C30A5.1 protein  
 C:Genetics:  
 A:Introns: 39/3; 67/2; 197/2; 264/3; 431/3  
 F:1-197/Product: hypothetical protein C30A5.1 #status predicted <CPR>

Query Match 5.8%; Score 186; DB 2; Length 489;  
 Best Local Similarity 21.9%; Pred. No. 1.2e-06;  
 Matches 108; Conservative 66; Mismatches 154; Indels 166; Gaps 19;

QY 50 SQVCRRWYEL-----DSLTR-----KHVTIALCVTTTPAR 79  
 DB 81 AQVCRSWILALDGSNWQVRDLFTFQDVKTAVVENLARCGGFKLSLKGCVNHDSA 140  
 QY 80 LR---RPPPHLESKLKGPRAAMFNLPEDWGGHVTPWKEISQYFCLKSLHFRMIV 136  
 DB 141 LRTFTSRCPNLEHLSL-----YRCKRV 162

140 VLS-----LAVGCKKLRRLHRLCIRNIVASVEMLEVFEGKLELELC-----180  
275 ISLPKALCRLGLTYIGKNEL-----PIVFMFAVLKLLDLYAMLDTEHCLMI 323  
181 -----IKNCRA-----IGEGDLIKLRNSWRKLTSLQFEVDANYRMYK-VYQDLOVERWPKQL 231  
324 QRCNPVLETRNVI--GDRGLEVLGRCKRLKRLRIERGDDDDQGDMEDEBEGTVSHRGLIA 381  
232 VPCDSLIVELSLGNCITAPGRGLACVLRNCKNLEKHL-----DMCTGVSDSD-----IIA 281  
382 LSQGCSELEVMAYVSDITNASLEHIGTHLKN.CDFRLVLLDHEEKITDLPDNGVRALL 441  
282 LVKASHLSISLRVPS-----DFTLPLANN-----ITLRLTDESLSAIA 321  
442 RGCDKLRFPALYLRG-----GLTDVGLGYIQYSPNVEMMLLVYGE--SDAGILLEFAK 494  
322 QHCKSLSPKISFSDGEFFSLFTLQGIITLQKCP-VRELSLDHVCVENDMG-MEALC 379  
495 GCPSLQKLEMRGCLFSEERALAAVATOLTSRLVLMV--QYGVSPSPGRDILLVMAPPFNIE 553  
380 SAQKLEILELVHCOEVSDEGL--ILVSQFSLNVLKSLKCLGVTDDG-----MRP-----427  
554 LIPSRKVATNTNPDETIVVE 573  
428 LVGSHKL-----ELLVVE 440

RESULT 12  
T08604  
hypothetical protein GRR1 - soybean  
C:Species: Glycine max (soybean)  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 11-Jun-1999  
C:Accession: T08604  
R:Chen, W.; Atherly, A.  
submitted to the EMBL Data Library, August 1997  
A:Reference number: Z15438  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Accession: T08604  
A:Molecule type: mRNA  
A:Residues: 1-690 <CHE>  
A:Cross-references: EMBL:AF019910; NID:g2407789; PID:g2407790  
A:Experimental source: variety L85-3044; root  
C:Genetics:  
A:Gene: grr1

Query Match 5.6%; Score 176.5; DB 2; Length 690;  
Best Local Similarity 20.6%; Pred. No. 1.1e-05;  
Matches 123; Conservative 76; Mismatches 170; Indels 227; Gaps 26;  
7 FLETLRS-----NMTEERNVRKTRVDVVLDC--VIPYIDDPKORDAVSQVRR 55  
81 FLFRGLASVFLHLSAENGLRSRSQRTSTESLPDECLFEILRLPAGQDRSVCASVSR 140  
56 WYELDSLTRKHVTIATCYTTTPARLRFRFPHLSLKLKGPRAAMFNLPEDNGHVTWP 115  
141 WMLLS-----SICKT-----EHSYSGTGNEN-----163  
116 VKEISQVFDCLSLHFRMIVKSDQLNLA--RDRGHVLHAKDKCS--GFTTDLGF 169  
164 -QEISDRGYLSRSLGKK--ATDVLAAIAGVTRASRG-LGKLTIRGNSDRGVTVNGLK 219  
170 HIGRFCKSLRVLFLEESLSLEKCEWHLALANNVTLETNFTYLTDAVKIELELLAK 229  
220 AIAHGCSFLKCSLNDVATVGDVG--LIEIASGCHQLEKLD--LCKCPNTSDKTLIAVAK 275  
230 NCPNLVSKLTDCEILDVNFVFKHASALEBFCGTYNEEPERYSAISLPAKLCRLGLTYI 289  
276 NCPNLALS-----284  
290 GKNELPVFMFAVLKLLDLYAMLDTEHCLMIQRCNLEVLETRNVIGRGLVLRGRC 349  
285 -----IESCPN-----IGNEGLQAICK-301

137 KSDQLONLARDRGHVLHAKDKCSGFTTDLGFHIGRFCKSLRVLFLESSILEKDGWL 196  
163 TDASCENLGR-YCHKLYNLNLENCSSITDRAMKYIGDGCNLSYLNISWCDALIDRGV--219  
197 HEALANN-TVLETFNY-----LTDIAVVKIELELLAKCNPVSVKLTDCETILDVNF 251  
220 -QILSNCKSLDILLRGCEGLTENVFGSVEAHMGAIKLNLOQLDITITQNIAN--276  
252 KHASALEEFCGGTYNEEPERYSAISLPAKLCRLGLTYIGKNELPIVFMFAVLKLLDLY 311  
277 -GATALEYLCWNCNQISDR-----SLVSLG-----301  
312 AMLDTEHCLMIQRCNLEVLETR--NVIGDRGLEVLGRCKRLKRLRIERGDDDDQGMED 369  
302 -----QHSHNLKVLSESGCTLLGDNQFIPLARGCQRLERLDM-----339  
370 BEGTVSHRGLIALSQGCSELEVMAY-VVSDITNASLEHIGT-HLKNLDCDFRLVLLDHEEK 427  
340 DCSLISDHTINSANNCTALRELSHCELITDESIGNLASKHRETL--NVLELDNCPQ 396  
428 ITDPLDNGVRALLRGCDKLRFPALY-----LRGGITDVGGLGYIQYSP-----472  
397 LTDSTLSH-----LHCKALKRIDLYDQNVSKAEIVRFQVVDVAVLSYDCIYFPCSTRS 451  
473 NVRWMLLVGVGESD 486  
452 TENWYKLTATKSD 465

RESULT 11  
G96837  
unknown protein T21f11.10 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: G96837  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luroso, J.S.; Maht, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: G96837  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-467 <STO>  
A:Cross-references: GB:AE005173; NID:g6730728; PIDN:AAF27118.1; GSPDB:GN00141  
C:Genetics:  
A:Gene: T21f11.10  
A:Map position: 1

Query Match 5.6%; Score 176.5; DB 2; Length 467;  
Best Local Similarity 23.6%; Pred. No. 6.4e-06;  
Matches 132; Conservative 69; Mismatches 186; Indels 171; Gaps 29;  
45 DRDAVSQVCRWYELDSLTRKHVTIATCYTTTPA-----RLRRPPLHLSLKLKGPRAA 99  
21 DRNLSLSCKRFFSLDNQVRSIRG-C-GLVPASDALLSLCRFPNLSKVEI-----71  
100 MENLIPDWGGHVTPWVKIEISQVFDCLSLHFRMIVKSDQLNLRDRGHVLHAKDK 159  
72 -----IYSGWMSKLGQVD-----DQGLLVLTN-CHSLTDLTSLF 106  
160 CSGFTTDLGFHIGRFCKSLRVLFLESSILEKDGWLHALLANNVTLETNFTYLTDA 217  
107 CTFITDVIGIHLSS-CPLESSL-----KLNFAPIITGG 139  
218 VKKIELELLAKCNPVSVKLTDC--BILDVNFVFKHASALEBFCGTYNEEPERYSA 274

QY	350	CKRLKRLRIER--GDDQGMDEBEGTVSHRGLI-ALSOQCSELEYMAVYVSDITNASLEH	406
DB	302	CENLRSISIKNGSGVGDQV-----AGLLSSASFALTQVKUESLTVSDLSLAVIGH	352
QY	407	IGTHLKNLDCDFRLVLLDHEEKITDPL-----DNG--VRALLRGCDKRLRFALYLRG	457
DB	353	YGV-----AVTDVLICLPNVSEKGFWMGNAHGLQKLTITINCCQ-	394
QY	458	GLTVDGLGYGVSPNVRMLL---GYGESDAGLLEFAKGCPSLOKLEMRGLFFSERA	514
DB	395	GVTVDGLEPIGRGCPNVQNLKRS AFL--SDKGLVSFAFAAPSVEBLSQ-----	441
QY	515	LAVATQLTSLRYLV-----QGVGPSGRDL--LVMAFPFWNIEL	554
DB	442	LAKSAHRTQIGLFGVFNCGAKLVLTILISYGIKDLNMDLPATSPSEISLTI	497
RESULT 13			
D96512			
hypothetical protein F2G19.16 [imported] - Arabidopsis thaliana			
C:Species: Arabidopsis thaliana (mouse-ear cress)			
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001			
C:Accession: D96512			
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alonso,			
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;			
ansen, N.F.; Hughes, B.; Huizar, L.			
Nature 408, 816-820, 2000			
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.			
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,			
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.			
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,			
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.			
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.			
A:Reference number: A86141; PMID:21016719; PMID:11130712			
A:Accession: D96512			
A:Status: preliminary			
A:Molecule type: DNA			
A:Residues: 1-518 <STO>			
A:Cross-references: GB:AE005173; NID:g11321778; PIDN:AG34255.1; GSPDB:GN00141			
C:Genetics:			
A:Gene: F2G19.16			
A:Map position: 1			
Query Match			
Best Local Similarity 5.4%; Score 172.5; DB 2; Length 518;			
Matches 107; Conservative 60; Mismatches 166; Indels 191; Gaps 20;			
QY	30	DVILDCVPIYIDDPKRDVAVSQVRRWYELDSLTKHVTI---ALCVTTTPARLRRFPFH	86
DB	45	DECLALVFQFENS-GNKRKALVCRRWMIYEGQNRVLSLHARSDLTISPS-LFSRFS	102
QY	87	LESILKLGKPRAMFNILPEDWGHVTPWVKEISQYFDCLKSLHFRMIVKDSLQNLAR	146
DB	103	VTKLSLKCDRRS-----VSIGDEALVKISLR-----CRNLKR	134
QY	147	DRGHVLHALKDKSGFTTDLPHIGRCKSLRVLF-----LEES	186
DB	135	-----LKLRACRELTDVGMFAFENCKDLKIFSCGCDFGAKGVKAVLDHCSNLEEL	186
QY	187	SILEKDEWLHELALNNVTLETLNFYLTDAVVKIEDLEIAKNCPLNVSKLFDCEILD	246
DB	187	SIKELRG-----FTDIA-----PEMIGPVA-----	207
QY	247	LVNFPFKHASALDEFC-GGTYNNEPERYSAISLPAKLCRLGLTYIGKNELPVIFMFAVLK	305
DB	208	-----ASSLKSICLKLNGQ-----C-FGPVIVGAKNKLKSLKLFRC-SG	245
QY	306	KLDLLYAMLTDEHCLM-----TORCPNLEVL-----ETRNIVGRGL	343
DB	246	DWDLLOQMSGKHGVVEIHLERMQVSDVALSAISYCSLSHLHLVKPTECTNF-----GL	301
QY	344	EVLGRCKKRLKRLRIERGGDDQGMDEBEGTVSHRGLIALSQSCSELEYMAVYVSDITNAS	403

DB	302	KAAAECKKRLKRLHIDGW-----KANLIGDEGLVAVAKFCSQLQELVLGVNPTTILS	353
QY	404	LEHIGTHLKNLDCDFRLVLLDHEEKITDPLDNGVRAALLRGCDKLRFRFALYLRGGGLTDVG	463
DB	354	LGMIAAKCLNL-----ERLALCGCDT-----FGDPE	379
QY	464	LCYIQOQYSPNVRMMLGYVGESDAGLLEFAKGCPSLOKLEMRGC	507
DB	380	LSCIAAKPALRKLCKNCPISDVGIENLANGCPGLTKVKIKKC	423
RESULT 14			
T41312			
probable DNA excision repair protein - fission yeast (Schizosaccharomyces pombe)			
C:Species: Schizosaccharomyces pombe			
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Jan-2000			
C:Accession: T41312; T41480			
R:Gwilliam, R.; Barrell, B.G.; Rajandream, M.A.; Wedler, H.; Wambutt, R.			
submitted to the EMBL Data Library, September 1998			
A:Reference number: Z21987			
A:Accession: T41312			
A:Status: preliminary; translated from GB/EMBL/DBJ			
A:Molecule type: DNA			
A:Residues: 1-563 <GWI>			
A:Cross-references: EMBL:AL031603; PIDN:CAA20907.1; GSPDB:GN00068; SPDB:SPCC330.02			
A:Experimental source: strain 972h; cosmid c330			
R:Bothe, G.; Pohl, T.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.			
submitted to the EMBL Data Library, September 1999			
A:Reference number: Z21997			
A:Accession: T41480			
A:Status: preliminary; translated from GB/EMBL/DBJ			
A:Molecule type: DNA			
A:Residues: 1-563 <BOT>			
A:Cross-references: EMBL:AL031644; PIDN:CAA21066.1; GSPDB:GN00068; SPDB:SPCC613.14			
A:Experimental source: strain 972h; cosmid c613			
C:Genetics:			
A:Gene: SPCC330.02; SPCC613.14			
A:Map position: 3			
A:Introns: 19/3			
Query Match			
Best Local Similarity 5.3%; Score 169.5; DB 2; Length 563;			
Matches 109; Conservative 76; Mismatches 158; Indels 199; Gaps 23;			
QY	71	LCVTTTPARLRRFPFHLESILKLGKPR-----AAMFNILPEDWGHVTP-----WKKEIS	120
DB	148	LCVPCSGRGAEDRSVPELRTKRKALTRKKVAAATMDEEIS-----VPKLQDL CIRVIA	200
QY	121	QYFDCLKSLHFRMIVKDSLQNLARDR-----GHVLHALKDKSGFTTGD	167
DB	201	EYINDIEAFGDIQGVNMNDKISQIIISKNSLNDTTVKILFLSGGQT--ELKLYDCSKITADS	258
QY	168	LPHIGRCKSLRVLFLE-----BSSILEKDEGMWLEHALN	202
DB	259	LFQIAYQCPNLQTLHTYTCQGMQDQVLFVADHHLTELTVSFQGAFLVSSSEWINFKKR	318
QY	203	NTVLETLNFYLTDAVVKIEDLEIAKNCPLNVSKLT-----DCEILDVNFPPKHASAL	257
DB	319	GSKLISLE--LTDARTHVSVINAVDCCPNLISLNSLRIFYLDDCEVRLLAGCRNLVSL	376
QY	258	E-EFCQGTYNNEPERYSAISLPAKLCRLGLTYIGKNELPVIFMFAVLKLDLLYAMLDT	316
DB	377	KIESPGGIIND-----GSTLDVNLQIGSGL--	401
QY	317	EDHCLMIQRCPNL--EVLERNVIGRGLGVIGRCKKRLKRLRIERGGDDQGMDE--DEE	371
DB	402	--HTLSLSCGCTKLTDEVLK-----QGIPGCCRLKHLNL-----SGLELTTDDE	443
QY	372	GTV-----SHRGLIALS-QGCSELEYMAVYVSDITNASLEHIGTHLKNLDCDFRLVLLDH	424
DB	444	ASIVFGEMKIQSGLETLSLRCLSLGDKTV-----RAVLVNS	480
QY	425	EKITDPLDNGVRALLRGCDKLRFRFALYLRGGGLTDVGLGYIQOYS-PNVNRMLLGVV-	482

Job time : 32 secs

Db 481 GHTLRITLDL-NM-----SFTDEALQIVNPPPLMLKALDVSWIR 520  
QY 483 GESDAGLLEFAKGCPSLOKLEMRGCLFFSERALAVAATQLTSLRYLVQGVGSPSGRDL 542  
Db 521 GMDKLVCDPESKPTLEKL-----LVWGDNHVLMPSNRLI 556  
QY 543 LV 544  
Db 557 LI 558

## RESULT 15

T45861  
Hypothetical protein F3A4.160 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Feb-2000  
C:Accession: T45861  
R:Barques, M.; Collado, M.C.; Navarro, P.; Terol, J.; Perez-Alonso, M.; Mewes, H.W.; May  
submitted to the Protein Sequence Database, December 1999  
A:Reference number: Z23007  
A:Accession: T45861  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-522 <BAR>  
A:Cross-references: EMBL:AL132978  
A:Experimental source: cultivar Columbia; BAC clone F3A4  
C:Genetics:  
A:Map position: 3  
A:Note: F3A4.160

Query Match 5.3%; Score 168.5; DB 2; Length 522;  
Best Local Similarity 21.9%; Pred. No. 3.2e-05;  
Matches 110; Conservative 53; Mismatches 160; Indels 179; Gaps 20;  
QY 45 DRDAVSQVCRWVELSLTRKHVTIALCYTTTPARLRFPFHLESILKXGKPRAAAFNLI 104  
Db 60 DRKRCSLVSKRWLLVDGQR-----HRLSLDAKSE----- 89  
QY 105 PEDWGHVTPWVKEISQYFDCLKSLHFR-----RMIVKSDDLONLARDRGHVLI 153  
Db 90 -----ILFLPFCIFNRFDSVTKALCRDRRSFSLSDALFIVSRCSNLIR----- 135  
QY 154 ALKLDKCGFTTDTGLPHIGFCCKSLRVL-----FLEESSILE-----KDGEM 195  
Db 136 -VALRCRRIITDGMESFARNCKSLKSCGCTFGAKGINAMLEHCKVIEELSLKIRG 194  
QY 196 LHELALNNTVLETNFIYT-DIAVVKIEDL-----ELLAKNCPNLVSVKLTDCBILD 247  
Db 195 LHELA-----EPKLSLSASLSRVFLKELVNGQVFGSLVATR--TLKKVKIIRC---L 242  
QY 248 VNFPPKASALEBFCGTYNEEBERYSAISLPKLCBLGTYIGKNELPIVFMFAAVLKKL 307  
Db 243 GNW-----DRVFMNGNGNSLATE-----IRLERLQVTDIG-----LFG----- 276  
QY 308 DLLYAMLDTEDHCLMFIQPCNLEVLERNV--IGDRGLEVLGRCKRLKRLRIERGDDQ 365  
Db 277 -----ISKCSNLETIHIVTPDCSNGLASVVERCKLRKHID----- 315  
QY 366 GMEDEBGTVSHRGLIALSQSCSELEYMAVVVSDITNASLEHIGTHLKNLCDFRLVLLDHE 425  
Db 316 GWRVKR--IGDQGLMSVAKHCLNLQELVLIGVDATYMSLS----- 353  
QY 426 EKITDPLDNGVALLRGCDKLRFPALYLRGGLTDVGLGYQGYSPNVVRWMLLGYVGES 485  
Db 354 -----AIASNCCKLERLAL-CGSGTIGDAEIGCIAEKCVCVTLRKFCIKGCLIS 399  
QY 486 DAGLLEFAKGCPSLQKLEMRGC 507  
Db 400 DVGVALALGCPKLVKLVKKC 421

Search completed: April 20, 2004, 01:38:54

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2004, 01:04:00 ; Search time 72 Seconds

(without alignments)  
2655.612 Million cell updates/sec

Title: US-10-009-791-22

Perfect score: 3180

Sequence: 1 TKTSAFLFTLSLRNMTEE.....QRSDPDTVVPLDTATCVDT 606

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2126.5	66.9	592	10 O04197	O04197 arabidopsis
2	1855.5	58.3	630	10 Q8W0G4	Q8W0G4 oryza sativ
3	1766.5	55.6	589	10 Q84QA7	Q84QA7 oryza sativ
4	1200	37.7	237	10 Q8H6H7	Q8H6H7 glycine max
5	819	25.8	585	10 Q9C5Y7	Q9C5Y7 arabidopsis
6	818	25.7	585	10 Q9ZRL2	Q9ZRL2 arabidopsis
7	813	25.6	585	10 Q94AU0	Q94AU0 arabidopsis
8	812.5	25.6	594	10 Q24660	Q24660 arabidopsis
9	760.5	23.9	575	10 Q9LW29	Q9LW29 arabidopsis
10	749.5	23.6	577	10 Q9LPW7	Q9LPW7 arabidopsis
11	741.5	23.3	619	10 Q9LTX2	Q9LTX2 arabidopsis
12	722.5	22.7	575	10 Q7XVM8	Q7XVM8 oryza sativ
13	720	22.6	635	10 Q9AUH6	Q9AUH6 populus tre
14	693	21.8	614	10 Q9STV5	Q9STV5 arabidopsis
15	693	21.8	623	10 Q8RWQ8	Q8RWQ8 arabidopsis
16	636.5	20.0	587	10 Q9FRB6	Q9FRB6 oryza sativ

17	621	19.5	603	10 Q8H7P5	Q8H7P5 oryza sativ
18	256.5	8.1	610	10 Q8LGK0	Q8LGK0 arabidopsis
19	249.5	7.8	610	10 Q9C5D2	Q9C5D2 arabidopsis
20	235	7.4	395	10 Q9SRR1	Q9SRR1 arabidopsis
21	229.5	7.2	618	10 Q9M004	Q9M004 arabidopsis
22	229	7.2	665	10 Q8RWU5	Q8RWU5 arabidopsis
23	221	6.9	642	10 Q8GZ31	Q8GZ31 arabidopsis
24	217	6.8	628	10 Q94GB3	Q94GB3 oryza sativ
25	217	6.8	628	10 Q7XGT5	Q7XGT5 oryza sativ
26	213	6.7	405	10 Q9FMW7	Q9FMW7 arabidopsis
27	213	6.7	772	5 Q9VF10	Q9VF10 drosophila
28	209.5	6.6	712	10 Q23399	Q23399 arabidopsis
29	203	6.4	154	10 Q82085	Q82085 lycopersico
30	193.5	6.1	2159	5 Q86IU5	Q86IU5 dictyosteli
31	188	5.9	601	10 Q9ARM5	Q9ARM5 arabidopsis
32	188	5.9	607	10 Q9ZWC6	Q9ZWC6 arabidopsis
33	187.5	5.9	628	10 Q9SKK0	Q9SKK0 arabidopsis
34	184	5.8	522	10 Q7X785	Q7X785 oryza sativ
35	183.5	5.8	466	5 Q8T3G0	Q8T3G0 caenorhabdi
36	180	5.7	479	10 Q7XV84	Q7XV84 oryza sativ
37	180	5.7	482	10 Q7XVD0	Q7XVD0 oryza sativ
38	179	5.6	436	4 Q961G2	Q961G2 homo sapien
39	176.5	5.6	467	10 Q9M8M4	Q9M8M4 arabidopsis
40	176.5	5.6	690	10 Q22512	Q22512 glycine max
41	175	5.5	422	11 Q9CZV8	Q9CZV8 mus musculu
42	175	5.5	423	4 Q9NVQ8	Q9NVQ8 homo sapien
43	175	5.5	423	11 Q8BHL6	Q8BHL6 mus musculu
44	172.5	5.4	440	10 Q7XVD2	Q7XVD2 oryza sativ
45	172.5	5.4	464	5 Q9V605	Q9V605 drosophila

#### ALIGNMENTS

RESULT 1

O04197 PRELIMINARY; PRT; 592 AA.  
 ID O04197  
 AC O04197;  
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein (At2G39940/T28M21.10).  
 GN T29M21.10 OR COI1  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RA Rounsley S.D., Ketchum K.A., Lin X., Phillips C.A., Brandon R.C.,  
 RA Fuhrmann J.L., White O., Kerlavage A.R., Adams M.D., Somerville C.R.,  
 RA Venter J.C.;  
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX PubMed=12244256;  
 RA Feys B.J., Benedetti C.S., Penfold C.N., Turner J.G.;  
 RL Plant Cell 6:751-759(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=98248619; PubMed=9582125;  
 RA Xie D.X., Feys B.F., James S., Nieto-Rostro M., Turner J.G.;  
 RT "COI1: an Arabidopsis gene required for jasmonate-regulated defense  
 and fertility";  
 RL Science 280:1091-1094(1998).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Kim C.J., Chen H., Cheuk R., Koesema E., Meyers M.C., Banh J.,  
 Bowser L., Carninci P., Dale J.M., Gibson H.A., Goldsmith A.D.,

RA Hayaishizaki Y., Ishida J., Jiang P.X., Jones T., Jones T., Kaniya A.,  
 RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu X.X.,  
 RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,  
 RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,  
 RA Tang C.C., Tortumi M., Yamada K., Yu G., Yu S., Shinozaki K.,  
 RA Davis R.W., Theologis A., Ecker J.R.;  
 RT "Arabidopsis cDNA clones";  
 RN Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RN SEQUENCE FROM N.A.  
 RP Cheuk R., Chen H., Kim C.J., Shinn P., Banh J., Bowser L.,  
 RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayaishizaki Y.,  
 RA Ishida J., Jones T., Kaniya A., Karlin-Neumann G., Kawai J., Lam B.,  
 RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,  
 RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,  
 RA Tang C.C., Tortumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,  
 RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;  
 RT "Arabidopsis ORF clones";  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF002109; AB95279.1; -;  
 DR EMBL; AF036340; AAC17498.1; -;  
 DR EMBL; AY045625; AAK73383.1; -;  
 DR EMBL; AY133556; AAM91386.1; -;  
 DR PIR; T52139; T52139.  
 SQ SEQUENCE 592 AA; 67665 MW; 1DDCF04990144C06 CRC64;

Query Match 66.9%; Score 2126.5; DB 10; Length 592;  
 Best Local Similarity 69.6%; Pred. No. 6.6e-173;  
 Matches 407; Conservative 72; Mismatches 97; Indels 9; Gaps 3;

QY 19 EERNVETRV-----VDVVLDCVPIYDDPKRDVAVSQVCRWYELDSLTKHVTALCY 73  
 Db 2 EDPDKRCKLSCVATVDDVDDVDDVDDVDDVDDVDDVDDVDDVDDVDDVDDVDDVDD 61  
 QY 74 TTPPAIRRRFPFLESILKKGKRAAMFNLIPEDDWGHVTPWVKEISQVDFCLKSLHFR 133  
 Db 62 TATPDLRRFPFNLRSILKKGKRAAMFNLIPEDDWGHVTPWVKEISQVDFCLKSLHFR 121  
 QY 134 MYKDSQLQNLARDGRHVLHALKDKCSGTTDGLFHIGRCKSLRVLFLEESSILEKDG 193  
 Db 122 MIVSDLDLRLAKARADDLTLKDKCSGTTDGLSLVTHCRKIKITLLMESSFSKDG 181  
 QY 194 EWLHELANNTVLETNFYLTADVVKIELELLAKNCPLVSVKLTDCIELDLVNFVKH 253  
 Db 182 KWLHELAQNTSLVLFNTFTEFAKISPKDLETIARNCRSLVSVKVGDFEILSELVFFKA 241  
 QY 254 ASALEFPCGTYNEE---PERYSAISLPALCRGLTYIGKNELPIVFMFAAVLKLDLL 310  
 Db 242 AANLEFPCGSLNEDIGMPEKYNLVFPRLCRGLSYMGPNEMPILFPFAAQIRKLDLL 301  
 QY 311 YAMLTEDHCKLTQPCNLEVLTRNVIGDRGLEVLGRCCRLKRLRIERGDDDDQWDE 370  
 Db 302 YALLETEDHCKLTQPCNLEVLTRNVIGDRGLEVLGRCCRLKRLRIERGDDDDQWDE 361  
 QY 371 EGVTVSHRGTLALSGCSELYMAVYSDITNASLHIGTHLKNLCPRLVLLDHEEKITD 430  
 Db 362 EGVTVSHRGTLALSGCSELYMAVYSDITNASLHIGTHLKNLCPRLVLLDHEEKITD 421  
 QY 431 LPLDNGVRLLRGCKLRLRFPALYLRGGLTDVGLGYIGQVSPNVRWMLLGVGESDAGLL 490  
 Db 422 LPLDNGVRLLRGCKLRLRFPALYLRGGLTDVGLGYIGQVSPNVRWMLLGVGESDAGLL 481  
 QY 491 EFAKGCPSLOKLEMRGCLFESERALVAATQTSRLRYLWVGQVSPSGRDLVWAPFW 550  
 Db 482 EFSRGCPNLQKLEMRGCC-FESRAIAAAVTKLPSRLYLMVQGRASMTGQDLQMARPYW 540  
 QY 551 NIELIPSRKVATNPDPETVVVEHPAHILAYVSLAGORSDFPDV 595  
 Db 541 NIELIPSRKVATNPDPETVVVEHPAHILAYVSLAGORSDFPDV 585

RESULT 2  
 Q8W0G4

Q8W0G4 PRELIMINARY; PRT; 630 AA.  
 AC Q8W0G4;  
 DT 01-MAR-2002 (trEMBLrel. 20, Created)  
 DT 01-MAR-2002 (trEMBLrel. 20, Last sequence update)  
 DE 01-MAR-2003 (trEMBLrel. 23, Last annotation update)  
 DE P0529E05.15 protein.  
 GN P0529E05.15.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]\_TaxID=4530;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC  
 RT clone:P0529E05.15";  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF003279; BAB84399.1; -;  
 DR Gramene; Q8W0G4; -;  
 DR InterPro; IPR001810; F-box.  
 DR InterPro; IPR001611; LRR.  
 DR Pfam; PF00646; F-box; 1.  
 DR Pfam; PF00560; LRR; 1.  
 SQ SEQUENCE 630 AA; 70619 MW; BFA12F8F1C93F23C CRC64;

Query Match 58.3%; Score 1855.5; DB 10; Length 630;  
 Best Local Similarity 60.9%; Pred. No. 1.1e-149;  
 Matches 352; Conservative 87; Mismatches 132; Indels 7; Gaps 5;

QY 28 VDVVLDCVPIYDDPKRDVAVSQVCRWYELDSLTKHVTALCYTTPPAIRRRFPFHL 87  
 Db 56 VPDEALHLVGHVEDPDEEAASVCRWRHRIADALTKHVTAFVCAAPARLRFPFRL 115  
 QY 88 ESILKKGKRAAMFNLIPEDDWGHVTPWVKEISQVDFCLKSLHFRMIKDSQLQNLARD 147  
 Db 116 ESLSLKGKRAAMFNLIPEDDWGHVTPWVKEISQVDFCLKSLHFRMIKDSQLQNLARD 175  
 QY 148 RGHVLHALKDKCSGTTDGLFHIGRCKSLRVLFLEESSILEKDGWHLHELANNTVLE 207  
 Db 176 RGHMLQELKDKCIGFTDGLRVARSCLRTLFEECHITDKGGEWLHELANNTVLE 235  
 QY 208 TLFNYLTADVVKIELELLAKNCPLVSVKLTDCIELDLVNFVKHSALEEFPCGTYNE 267  
 Db 236 TLFNYLTADVVKIELELLAKNCPLVSVKLTDCIELDLVNFVKHSALEEFPCGTYNE 294  
 QY 268 --EPERYSAISLPALCRGLTYIGKNELPIVFMFAAVLKLDLLYAMLTEDHCKMLIQ 325  
 Db 295 VGELTYKVKVFPFRLCFGLTYMGNTNMPVFPFMSKLLKLDLQYTLTDEHCKLIK 354  
 QY 326 CNLNVLETRNVIGDRGLEVLGRCCRLKRLRIERGDDDDQWDEBEGTVSHRGTLALSQ 385  
 Db 355 CPNLLILEVRNVIGDRGLEVLGRCCRLKRLRIERGDDDDQWDEBEGTVSHRGTLALSQ 414  
 QY 386 CSELEYMAVYSDITNASLHIGTHLKNLCPRLVLLDHEEKITDPLONGVRLLRGCD 445  
 Db 415 CRELEYIAAVYSDITNASLHIGTHLKNLCPRLVLLDHEEKITDPLONGVRLLRGCD 474  
 QY 446 KLRRFPALYLRGCKLRLRFPALYLRGGLTDVGLGYIGQVSPNVRWMLLGVGESDAGLL 505  
 Db 475 KLRRFPALYLRGCKLRLRFPALYLRGGLTDVGLGYIGQVSPNVRWMLLGVGESDAGLL 534  
 QY 506 GCLFFESERALVAATQTSRLRYLWVGQVSPSGRDLVWAPFWNIELIPSRKVATN-- 563  
 Db 535 SCC-FESERALVAATQTSRLRYLWVGQVSPSGRDLVWAPFWNIELIPSRKVATN-- 593  
 QY 564 TNPDETVVVEHPAHILAYVSLAGORSDFPDV 601  
 Db 594 TEDGEPVDSH-AQVLAAYVSLAGORSDFPDV 630

RESULT 3



DT	01-WAR-2003 (TReMBLrel. 23, Last sequence update)	
DE	01-WAR-2003 (TReMBLrel. 23, Last annotation update)	
DE	Putative coronatine-insensitive 1 (Fragment).	
GN	COIL1	
GN	Glycine max (Soybean)	
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;	
OC	eurosid I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.	
OX	NCBI_TaxID=3847;	
OX	[1]_RN	
RP	SEQUENCE FROM N.A.	
RP	STRAIN=cv. TGX1835-2E;	
RC	Christiansen L.C.; Ulvskov P.;	
RC	"Putative Glycine max sBOC11 homologous to Arabidopsis COI (atFBL2)." ;	
RL	Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.	
RL	EMBL; AF536527; AAN31713.1; -.	
FT	NON TER 1	
FT	NON TER 237	
FT	SEQUENCE 237 AA; 26280 MW; 6P8C6B17AF74E1FD CRC64;	
SQ		
Query Match 37.7%; Score 1200; DB 10; Length 237;		
Best Local Similarity 99.1%; Pred.No. 2.9e-94;		
Matches 230; Conservative 1; Mismatches 1; Indels 0; Gaps		
QY	374 VSHRGHIALSQCSSELYMAVVVSDDITNASLEHIGTHLKNLCDPRLVLLDHEEKITDLP	
DB	6 VSHRGHIALSQCSSELYMAVVVSDDITNASLEHIGTHLKNLCDPRLVLLDHEEKITDLP	
QY	434 DNGVPAIIARGCDKLRFAIYLRGGITDVGIGYGYSPNVRWMLLYGVGSDAGLLPFA	
DB	66 DNGVPAIIARGCDKLRFAIYLRGGITDVGIGYGYSPNVRWMLLYGVGSDAGLLPFA	
QY	494 KGCPSLQKLEMRGCIFFSERALAVAAATQLTSLRYLWVGYGVSFGRDLLVWAPFFWNIE	
DB	126 KGCPSLQKLEMRGCIFFSERALAVAAATQLTSLRYLWVGYGVSFGRDLLVWAPFFWNIE	
QY	554 LIPSRKVATNPDPETVVVEHPAIIAYSLAGORSDDPDTVPVLDATATCVD 605	
DB	186 LIPSRKVATNPDPETVVVEHPAIIAYSLAGORSDDPDTVPVLDATATMLD 237	
RESULT 5		
Q9CSY7		
ID	Q9CSY7 PRELIMINARY; PRT; 585 AA.	
AC	Q9CSY7;	
DT	01-JUN-2001 (TReMBLrel. 17, Created)	
DT	01-JUN-2001 (TReMBLrel. 17, Last sequence update)	
DE	01-JUN-2003 (TReMBLrel. 24, Last annotation update)	
DE	GRR1-like protein 1.	
GN	GRH1	
GN	Arabidopsis thaliana (Mouse-ear cress).	
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;	
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsis.	
OX	NCBI_TaxID=3702;	
OX	[1]_RN	
RP	SEQUENCE FROM N.A.	
RP	Thelander M., Fredriksson D., Schouten J., Hoge H.C., Ronne H.;	
RT	"Cloning by pathway activation in yeast: identification of an	
RT	Arabidopsis thaliana F-box protein that turns on glucose repression."	
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.	
RL	EMBL; AF291816; AAK01147.1; -.	
DR	InterPro: IPR001810; F-box.	
DR	InterPro: IPR001511; LRR.	
DR	Pfam: PF00646; F-box; 1.	
DR	Pfam: PF00560; LRR; 1.	
DR	SMART; SMO0256; FBOX; 1.	
SQ	SEQUENCE 585 AA; 65678 MW; CAF4E7C2B4150BDD CRC64;	

SQ SEQUENCE 385 AA; 65678 MW; CAI4E/CZB4150BDD CRC64;  
 Query Match 25.8%; Score 819; DB 10; Length 585;  
 Best Local Similarity 34.6%; Pred. No. 4e-61;  
 Matches 200; Conservative 109; Mismatches 237; Indels 32; Gaps

```

QY 32 VLDGVIPYIDDPKORDAVSQVCRWYELDSLTRKHVTIALCYTTTPARLRERRRPHLESILK 91
D 9 VLEHILSPIDSNEDRNSVSLCKSWFETERKTRKRVFVGNVCYAVSPAATRRFPPEMSRLT 68
QY 92 LKGPRAAMENLIPEDGHHVTPWVKELSQYDFCLKSLHFRMIVKUSDIONLA---RDR 148
D 69 LKGPHFADYNLVPDGGGYAWPWEIAWAKSSLSLEIRMKRMVVTDECLEKIAAFKD- 127
QY 149 GHVLHAIKDKCSGFTTDLGPHIGRCKSLRVLFLRESSILEKDGEMHLEALANTVLET 208
D 128 ---FKVLVTSCGFTSDGIAAIASTCRNLRLVLELRECEIVEDLGDMLSYFPESSTSLVS 184
QY 209 LNFYLTDAVVKIEDLELLAKNCNLVSVKLTDCETILD-INVFFKHASALEEFCGGTYNE 267
D 185 LDFSCLD-SEVKISDLERLVSRSNLSKLNPAVTLGLVSLRLCAPQTELGTGSPFAA 243
QY 268 E--PERYSALSIPAKLCK----LGLTVIGKNELPIVFMFAAVLKKLDLLYAMLDTHC 320
D 244 QLKPEAFSKLSEAFNSCKQLSGLMDVLPYLPALYSVCPGLTSLNTYATVRMPDLV 303
QY 321 MLQRCNLVLETRNVIGDRGLVGRCKRLKRLRIERGDDDDQMGDEEGTVSHRGLI 380
D 304 ELLRRCCKLQKLVMDLIEDKLEAVASYCKELRLAVFPSEPD--LDATNIPLEQGLV 361
QY 381 ALSQGCSELEVMAYVSDITNASLEHIGTHLKNLCDFRLVLLD--HEEKITDPLDNGVR 438
D 362 FVSKGCKRLSVLFCVQFTNAALFTIARKRNLCFRLCVIEPAPDYKTNFPLDKGFK 421
QY 439 LQKLEMRGCLFFERLALAVATQTLRLVYVQGVSPSGRDLVLMARFPFNIPLPSR 558
D 479 LKLEIRDCP-FGDTALLEHAAKLETMSLWMSSCFVSPGACKLISQKMPRLNVEVI--- 534
QY 559 KVATNTPDETIVVVEHPA-HILAYISLAGORSDFPDVT 595
D 535 ----DEHPESPRESSPVERIYRTVAGPRMDTPEFV 568

```

## RESULT 6

```

Q9ZR12 ID Q9ZR12 PRELIMINARY; PRT; 585 AA.
AC Q9ZR12;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Putative homolog of transport inhibitor response 1 (Putative F-box protein AtPBL18)
DE F4C21.11 OR AT4G03190.
GN Arabidopsis thaliana (Mouse-ear cress).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]_TaxID=3702;
RN Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RA Habermann K., de la Bastide M., Huang E.N., Gnoj L., Schutz K.,
RA Preston R., Calma C., Martienssen R., Parnell L.D., Dedhia N.,
RA McCombie W.R.;
RT "Arabidopsis thaliana BAC F4C21 from chromosome IV near 17 cM.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN
RN SEQUENCE FROM N.A.
RA Spiegel L.A., Huang E.N., Nascimben L.U., de la Bastide M., Vil D.M.,
RA Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,
RA Shekher M., Schutz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,
RA Mewes H.W., Lemcke K., Meyer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RP EU Arabidopsis sequencing project;
RA

```

Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

## SEQUENCE FROM N.A.

Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan V.W., Lee J.M.,  
Quach H.L., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,  
Yu G., Yuan S., Chen H., Cheuk R., Jones T., Kim C.J., Nguyen M.,  
Palm C.J., Shinn P., Southwick A., Tripp M.G., Wu T., Davis R.W.,  
Ecker J.R., Theologis A.,  
"Arabidopsis Open Reading Frame (ORF) Clones";  
Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

EMBL; AC005275; AAD14447.1; -

EMBL; AL161496; CAB77804.1; -

EMBL; AY150427; AAN12969.1; -

PIR; E85040; E85040.

InterPro; IPR001810; F-box.

Pfam; PF00646; F-box; 1.

Pfam; PF00560; LRR; 1.

SMART; SM00256; FBOX; 1.

SEQUENCE 585 AA; 65647 MW; D39D627C82864D83 CRC64;

Query Match 25.7%; Score 818; DB 10; Length 585;

Best Local Similarity 34.6%; Pred. No. 4.8e-61;

Matches 200; Conservative 109; Mismatches 237; Indels 32; Gaps 12;

QY 32 VLDGVIPYIDDPKORDAVSQVCRWYELDSLTRKHVTIALCYTTTPARLRERRRPHLESILK 91

D 9 VLEHILSPIDSNEDRNSVSLCKSWFETERKTRKRVFVGNVCYAVSPAATRRFPPEMSRLT 68

QY 92 LKGPRAAMENLIPEDGHHVTPWVKELSQYDFCLKSLHFRMIVKUSDIONLA---RDR 148

D 69 LKGPHFADYNLVPDGGGYAWPWEIAWAKSSLSLEIRMKRMVVTDECLEKIAAFKD- 127

QY 149 GHVLHAIKDKCSGFTTDLGPHIGRCKSLRVLFLRESSILEKDGEMHLEALANTVLET 208

D 128 ---FKVLVTSCGFTSDGIAAIASTCRNLRLVLELRECEIVEDLGDMLSYFPESSTSLVS 184

QY 209 LNFYLTDAVVKIEDLELLAKNCNLVSVKLTDCETILD-INVFFKHASALEEFCGGTYNE 267

D 185 LDFSCLD-SEVKISDLERLVSRSNLSKLNPAVTLGLVSLRLCAPQTELGTGSPFAA 243

QY 268 E--PERYSALSIPAKLCK----LGLTVIGKNELPIVFMFAAVLKKLDLLYAMLDTHC 320

D 244 QLKPEAFSKLSEAFNSCKQLSGLMDVLPYLPALYSVCPGLTSLNTYATVRMPDLV 303

QY 321 MLQRCNLVLETRNVIGDRGLVGRCKRLKRLRIERGDDDDQMGDEEGTVSHRGLI 380

D 304 ELLRRCCKLQKLVMDLIEDKLEAVASYCKELRLAVFPSEPD--LDATNIPLEQGLV 361

QY 381 ALSQGCSELEVMAYVSDITNASLEHIGTHLKNLCDFRLVLLD--HEEKITDPLDNGVR 438

D 362 FVSKGCKRLSVLFCVQFTNAALFTIARKRNLCFRLCVIEPAPDYKTNFPLDKGFK 421

QY 439 LQKLEMRGCLFFERLALAVATQTLRLVYVQGVSPSGRDLVLMARFPFNIPLPSR 558

D 479 LKLEIRDCP-FGDTALLEHAAKLETMSLWMSSCFVSPGACKLISQKMPRLNVEVI--- 534

QY 559 KVATNTPDETIVVVEHPA-HILAYISLAGORSDFPDVT 595

D 535 ----DEHPESPRESSPVERIYRTVAGPRMDTPEFV 568

RESULT 7

Q94AU0 ID Q94AU0 PRELIMINARY; PRT; 585 AA.

AC Q94AU0; (Tremblrel. 19, Created)

DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)





```

RN RP SEQUENCE FROM N.A.
RA Nguyen M., Karlin-Neumann G., Southwick A., Tripp M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC012187; AAF78487.1; -
DR EMBL; AF099541; AAM20393.1; -
DR EMBL; BT002118; AAN72129.1; -
DR PIR; F86261; F86261.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR008945; Skp1_Skp2.
DR Pfam; PF00646; F-box; 1.
SQ SEQUENCE 577 AA; 64906 MW; 8E419B4D80368661 CRC64;

Query Match      23.6%; Score 749.5; DB 10; Length 577;
Best Local Similarity 33.8%; Pred. No. 3.4e-55;
Matches 197; Conservative 97; Mismatches 258; Indels 31; Gaps 13;

QY 30 DVVLCVPIYDDPKDRDAVSQVCRRWYELDSLTKHVTIALCYTTTPARLRRFPFHL 89
Db 6 DEVIEHVDFVASHKDRNSISLVCKSWHKIERFSRKEVFIGNVCYAINPERLIRFPCLKS 65
QY 90 LKLGKPRAMNLIPEDMGHHVTPWVKELISQYFDCLSLHFRMTVKDSLDLQNLARDG 149
Db 66 LTLGKPHFADPNLVPHMGVGVHFWIEALARSVGLBELRLKRWVVTVDESLLLSRFA 125
QY 150 HVLHAKLDKCSGFTDGLFHIGRCKSLRVLFLEESSILEKDGWHLHNLNNTVLETL 209
Db 126 N-FKSLVLSVCSGFTDGLASIAANCRLHRELDLQNEIDHRRGWLNCFFPSCCTLLMSL 184
QY 210 NFYL-----TDIAVKIEDLELLAKNCPNLVSVKLTDCIELD-LVNFPHKASALEEFCOGT 264
Db 185 NFACLGKGETNVA-----LERLVARSPLNLSKLNRAVPLDALARLMSCAPQLVDLGVGS 239
QY 265 YNEEPERYSAISLPKLCRL-----GLTYGKNEPLIVFVFAVLKLLDLYLMLDTE 317
Db 240 YENEDPESFAKMTAKYTSLSRSLSGFLEVAPICLPFAFYPICQNLISLNLISYAAEIQG 299
QY 318 DHCM-LIQCPLNLEVLRNVIGDRGLGVIGRCCRLKRLRIERGGDDQGMEDSGTVSH 376
Db 300 NHLIKLIQCKLQRLWILDSIGDKGLAVVAATCKELQELRV-FPSDVHGEEDNNAVTE 358
QY 377 RGLIALSGCSELEFVWYVSDITWASIEHIGTHLKNLDFRLVLDD-HE-EKITDLPID 434
Db 359 VGLVAISGCPKLHSLYFCQMTNAALIAVAKNCNPFIRFLCTILEPHKPDHITFQSLD 418
QY 435 NGVALLRGCKLRRLFALYLRGGTLTDVGLGVIGYQVSPNRMWMLLGYVGSDEAGLLEPAK 494
Db 419 EGFGAIVACKGLRLSV---SGLTDQVFLYIGMAEQLEMLSTAFAGDITKGMVLVIN 475
QY 495 GCPSLQKLEMRGCLFFSERALAAVATQTSRLYLWVGYSVPSGRDLLVMARPPWNIEL 554
Db 476 GCKMKRLEIRDSP-FGNAALLADGVRYETMRSLWMSCEVTGLGCKELAQNSPRLNVEI 534
QY 555 IPSRK--VATNTPDETIVVVEHPAHILAYSLAGORSFPDPTV 595
Db 535 INENENNGMEQNEEREKVD---KLTYLRYTVVGTGRDAPPV 574

RESULT 11
Q9LTX2 PRELIMINARY; PRT; 619 AA.
AC Q9LTX2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Transport inhibitor response 1 protein (At5g49980/k9p8_12).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

```

OC OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OX eurosids II; Brassicales; Brassicaceae; Arabidopsis.
RN [1] NCBI_TaxID=3702;
RP SEQUENCE FROM N.A.
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT clones.";
RN DNA Res. 7:31-63(2000).
RL [2]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cdna clones.";
RN Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RL [3]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Bowser L., Carninci P.,
RA Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y.,
RA Hsuan V.W., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M.,
RA Southwick A., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,
RA Yamada K., Yu G., Yuan S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cdna clones.";
RN Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RL [4]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
RA Chan M.M., Chang C.H., Dale J.M., Hayashizaki Y., Hsuan V.W.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G., Yuan S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis ORF clones.";
RN Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; AB024C32; BAA97019.1; -
DR EMBL; AY056431; AAL08287.1; -
DR EMBL; AY139774; AAM98092.1; -
DR EMBL; BT004536; AAO42782.1; -
DR InterPro; IPR001810; F-box; 1.
DR Pfam; PF00646; F-box; 1.
SQ SEQUENCE 619 AA; 69316 MW; D0614AF071EE4FD2 CRC64;

Query Match      23.3%; Score 741.5; DB 10; Length 619;
Best Local Similarity 33.4%; Pred. No. 1.8e-54;
Matches 196; Conservative 100; Mismatches 255; Indels 35; Gaps 12;

QY 28 VDVVLCVPIYDDPKDRDAVSQVCRRWYELDSLTKHVTIALCYTTTPARLRRFPFHL 87
Db 54 VLENVLENVLQFLDSRCDRNAASLVCKSWHVRVEALTRSEVFIGNVCYALSPARLTQRFKRV 113
QY 88 ESKLKGKPRAMNLIPEDMGHHVTPWVKELISQYFDCLSLHFRMTVKDSLDLQNLARD 147
Db 114 RSLVLKGRPADFNLMPPDMGANFAPWVSTMAQAYPCLEKVDLKRMEVTDLDLALLA-D 172
QY 148 RGHVLHAKLDKCSGFTDGLFHIGRCKSLRVLFLEESSILEKDGWHLHNLNNTVLE 207
Db 173 SFGFKELILVCCSGFGTSGISIVANKRKLKVLIDLESEVTDDEVDVMSIFPEDVTCL 232
QY 208 TLFYLTIDIAVKIEDLELLAKNCPNLVSVKLTDC-CEILDVNFPHKASALEEFCOGTY- 265

```





Search completed: April 20, 2004, 01:38:18  
Job time : 75 secs





Creation date: 06-03-2004  
Indexing Officer: SUK - SOPHY UK  
Team: OIPEBackFileIndexing  
Dossier: 10009791

Legal Date: 05-17-2004

No.	Doccode	Number of pages
1	SRNT	2

Total number of pages: 2

Remarks:

Order of re-scan issued on .....